BMJ Open

External Validation of the 4C Mortality Score for Hospitalized COVID-19 Patients in the RECOVER network

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2021-054700
Article Type:	Original research
Date Submitted by the Author:	10-Jul-2021
Complete List of Authors:	Gordon, Alexandra June; Stanford University School of Medicine, Emergency Medicine Govindarajan, Prasanthi; Stanford University School of Medicine, Department of Emergency Medicine Bennett, Christopher L.; Stanford University School of Medicine, Department of Emergency Medicine Matheson, Loretta ; Stanford University School of Medicine, Department of Emergency Medicine Kohn, Michael; Stanford University School of Medicine, Department of Emergency Medicine Camargo, Carlos; Massachusetts General Hospital, Harvard Medical School, Emergency Medicine Kline, Jeffrey ; Indiana University School of Medicine, Emergency
Keywords:	COVID-19, ACCIDENT & EMERGENCY MEDICINE, Adult intensive & critical care < ANAESTHETICS, EPIDEMIOLOGY, GENERAL MEDICINE (see Internal Medicine)

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

External Validation of the 4C Mortality Score for Hospitalized COVID-19 Patients in the RECOVER network

Alexandra June Gordon, MD (First author)
Department of Emergency Medicine, Stanford University School of Medicine,
ajgordon@stanford.edu
Prasanthi Govindarajan, MBBS, MAS (Co-first author)
Department of Emergency Medicine, Stanford University School of Medicine
prashagovindarajan@stanford.edu
Christopher L. Bennett, MD
Department of Emergency Medicine, Stanford University School of Medicine
cleebennett@stanford.edu
Loretta Matheson, MS
Department of Emergency Medicine, Stanford University School of Medicine
lorettamatheson@stanford.edu
Michael A. Kohn, MD
Department of Emergency Medicine, Stanford University School of Medicine
Michael.Kohn@ucsf.edu
Carlos A. Camargo, Jr., MD, DrPH
Department of Emergency Medicine, Massachusetts General Hospital, Harvard Medical School, Boston, MA
CCAMARGO@PARTNERS.ORG
Jeffrey A. Kline MD
Department of Emergency Medicine, Indiana University School of Medicine
jefkline@iu.edu

Corresponding author: Jeffrey Kline MD, jefkline@iu.edu
Department of Emergency Medicine,
Indiana University School of Medicine
5th/3rd Bank Building, 3rd Floor
640 Eskenazi Avenue
Indianapolis, IN 46202
Telephone: 317-670-0541

Manuscript word count (not including title, abstract, acknowledgment, references, tables, and figure legends): 2,450

ABSTRACT:

Objectives: Estimating mortality risk in hospitalized SARS-CoV-2+ patients may help with choosing level of care and discussions with patients. The Coronavirus Clinical Characterisation Consortium Mortality Score (4C Score) is the most promising of several published COVID-19 mortality risk models. We examined the association of patient-level risk factors with 30-day mortality in hospitalized SARS-CoV-2+ patients and investigated the discrimination and calibration of the 4C Score. This was a retrospective cohort study of SARS-CoV-2+ hospitalized patients within the RECOVER (REgistry of suspected COVID-19 in EmeRgency care) Network.

Setting: 99 emergency departments (EDs) across the US.

Participants: Patients ≥ 18 years old who tested positive for SARS-CoV-2 in the ED and were hospitalized.

Primary outcome: Death within 30 days of the index visit. We performed logistic regression analysis, reporting multivariable risk ratios (MVRRs). We also calculated the AUROC and mean prediction error for the original 4C Score and after dropping the c-reactive protein component(CRP).

Results: We included 7,961 hospitalized COVID-19 patients of whom 1847 (23.2%) died within 30 days. The 30-day mortality was increased with age 80+ years (MVRR = 4.92, 95% CI 3.81-6.03); male sex (MVRR = 1.11,1.03-1.20); and nursing home/assisted living facility residence (MVRR = 1.37, 1.24-1.51). The 4C Score had comparable discrimination in the RECOVER dataset compared with the original 4C validation dataset (AUROC: RECOVER 0.779 (95% CI 0.768-0.790), 4C validation 0.763 (95% CI 0.757-0.769). Score-specific mortalities in our sample were lower than in the 4C validation sample (mean prediction error 2.5%). Dropping the CRP component from the 4C Score did not substantially affect discrimination but would underestimate risk (mean prediction error -3.3%).

Conclusions: We independently validated 4C Score as predicting risk of 30-day mortality in hospitalized SARS-CoV-2+ patients. We recommend dropping the CRP component of the score and using our recalibrated mortality risk estimates.

ARTICLE SUMMARY:

Strengths and Limitations:

- In this first study using a national sample of patients who tested positive for SARS Co-V 2 and hospitalized through the emergency departments, our results confirmed the previous findings that including older age, comorbidities, BMI ≥ 40 kg/m2, higher respiratory rate, and lower oxygen saturation were associated with 30-day mortality.

-We also observed that the arrival to the emergency care setting from a nursing home was associated with increased mortality.

-We independently validated 4C Mortality Score as predicting risk of 30-day mortality in hospitalized SARS-CoV-2+ patients.

-We recommend dropping the CRP component of the score and using our recalibrated mortality risk estimates when estimating the 30-day mortality in hospitalized patients who test positive for SARS CoV2.

INTRODUCTION

The COVID-19 pandemic has placed tremendous strain on emergency and critical care resources in hospitals worldwide.(1-3) To prepare the healthcare systems for the surges, several studies have developed prediction models to assess mortality risk in patients hospitalized with COVID-19. These studies identified the following risk factors for mortality or critical care admission: age, sex, comorbid conditions, laboratory values, and vital signs.(4-16)

In a systematic review that evaluated many of these risk prediction studies using the prediction model risk of bias assessment tool (PROBAST), Wynants et al. concluded that many of the current risk models may be misleading.(10) However, the authors' analysis suggested that one COVID-19 mortality prediction model, the 4C Mortality Score, which was built on a large UK data set, had relatively low risk of bias in most domains by the PROBAST criteria. The 4C Mortality Score includes eight variables - age, sex, number of comorbidities, respiratory rate, oxygen saturation, level of consciousness, blood urea nitrogen, and C-reactive protein (CRP).(17)

Table 1. Point assignment for 4C Mortality Score

Age Group (years)	**18 - 49**	
	50 - 59	+ 2
	60 - 69	+ 4
	70 - 79	+ 6
	80+	+ 7
Sex at birth	**Female**	
	Male	+ 1
Comorbidities	**0**	
	1	+ 1
	2 +	+ 2
Respiratory rate (breaths/minute)	**< 20**	
	20 - 29	+ 1
	≥ 30	+ 2
Oxygen saturation, room air (%)	**≥ 92**	
	< 92	+ 2
Altered mental status*	**No**	
	Yes	+ 2
BUN (mg/dL)†	**< 20**	
	20 - 39	+ 1
	≥ 40	+ 3

C-reactive protein (mg/dL)†	**< 5.0**	
	5.0 - 9.9	+ 1
	≥ 10.0	+ 2

* Altered mental status is patient-reported symptom 4C used Glasgow coma scale < 15.
† BUN and C-reactive protein units converted from 4C.

In this study, we investigated the risk of 30-day mortality in hospitalized SARS-CoV-2+ patients within the RECOVER (REgistry of suspected COVID-19 in EmeRgency care) Network.(18) In a large cohort of SARS-CoV-2+ patients hospitalized from 99 US emergency departments (EDs), we determined the relation of demographic and clinical factors with 30-day mortality and investigated the discrimination and calibration of the 4C Mortality Score with and without the CRP value.

METHODS

In this retrospective cohort study, we included patient-level data from the RECOVER Network, a national registry of patients who were tested for SARS-CoV-2 during their ED visit. We restricted the analysis to patients ≥ 18 years old who tested positive for SARS-CoV-2 and were hospitalized from the ED.(18) The study was approved or deemed exempted as non-human subject research by the Institutional Review Boards of all participating sites.

Data Source

We obtained data from 40 medical centers representing 99 EDs from 27 US states and the District of Columbia. Data were collected using a REDCap data collection form that was distributed to the ED sites during the study period (3/2020-9/2020); our data were downloaded from the registry in 12/2020. The REDCap form (Appendix A) had seven sections and 204 questions, which generated 360 data elements. Variables reflected a combination of routinely collected information (e.g. patient demographics, medical history, vital signs, and diagnostic test results), patient-reported symptoms and risk exposures, clinical outcomes (e.g. admission, therapies, death), and those deemed important by the RECOVER Network steering committee. After creation, but prior to launch, the data form was piloted at 19 sites and refined. For additional section details and the questions, please refer to the data collection form in the appendix (supplement). The data were obtained from the electronic healthcare record using a combination of electronic download for routinely collected, coded variables (e.g., age, vital signs and laboratory values), supplemented by chart review by research personnel, using methods previously described.(18)

Patient Involvement

Patients were not involved in the development of our work, setting the research question or determining the outcome measures. This applies to both the RECOVER Network and the work presented here. Given the nature and limitations of emergency care during the COVID-19 pandemic, it was not appropriate or possible to involve patients or the public in the design, or conduct, reporting, or dissemination plans of our work.

Study Variables

We analyzed patient characteristics such as demographics, vital signs, symptoms, risks for infection, comorbidities, and medications. Following the 4C Mortality Score, we categorized the patients into five age groups (18-49, 50-59, 60-69, 70-79, 80+). The US standard ethnicity (Latinx/Hispanic yes/no) and race categories were combined into 8 categories (Non-Hispanic White, Non-Hispanic African American/Black, Non-Hispanic Asian, Non-Hispanic Native American/Alaskan Native, Non-Hispanic Pacific Islander/Native Hawaiian, Non-Hispanic mixed, and Non-Hispanic Unknown). In the analysis, we combined Non-Hispanic Native American (0.2%), Non-Hispanic Pacific Islander (0.2%), Non-Hispanic mixed, Other, and Unknown into a single group (12%).

All included patients had a positive reverse transcriptase polymerase chain reaction test (RT-PCR) test for SARS-CoV-2. Almost all the tests were performed during the ED visit, but we also included patients who had a test in a physician's office or urgent care immediately prior to the ED visit. We excluded patients whose 30-day vital status could not be ascertained and those who died in the ED before vital signs were recorded.

Study Outcome

The primary outcome was proportion of hospitalized SARS-CoV-2+ patients who died within 30 days of the index visit and the 4C mortality score's predictive accuracy as measured by the area under the ROC curve (AUROC) and mean prediction error.

Data Preparation and Statistical Analysis

For comparison with the other cohorts, we report the median and interquartile range of continuous variables – both in the entire cohort and in the subgroup who died within 30 days – and compared the median values using the rank sum test. We performed univariable analysis on 25 independent variables that were included on the data collection form using complete (non-missing) data and reporting risk ratios for 30-day mortality. For risk ratio reporting of continuous variables, we chose category boundaries based

on the 4C Mortality Score (age, respiratory rate, oxygen saturation, BUN, and CRP) or other published mortality prediction models (BMI, creatinine, total bilirubin).(19)

We selected variables for our multivariable logistic regression model based on the 4C Mortality Score, other prior studies, and clinical judgement. The RECOVER dataset has complete data (< 3% missing) on most variables, with the exception of CRP, smoking status, body mass index (BMI), and bilirubin. For our multivariable analysis, we used imputed values for missing data using Stata's implementation of the data augmentation (DA) algorithm.(20) We report multivariable risk ratios with 95% confidence intervals. Statistical analysis was performed using SAS Enterprise Guide Version 8.3 and Stata/SE 16.1.(21)

We replicated the 4C Mortality Score described by Knight et al with one modification. Since we did not have a variable for Glasgow Coma Score or confusion on exam, we used the symptom "altered mental status or confusion" instead. In addition to the full score, we tested a modified score dropping CRP, which was missing in 40% of the records. We evaluated discrimination and calibration using 9 score categories available from Knight et al: (0-2, 3-4, 5-6, 7-8, 9-10, 11-12, 13-14, 15-16, ≥17). We used the mortality reported in the 4C validation dataset as our predicted risks for comparison with observed mortality. For reporting, we pooled the results into the 4 risk groups defined by Knight et al: Low 0-3, Intermediate 4-8, High 9-14, and Very High ≥15. The AUROC was calculated with 95% confidence interval using the DeLong method.(22) Calibration was assessed using a standard calibration table, mean prediction error, and the square root of both the calibration error and the Brier score. We also used a modified Bland-Altman-style calibration plot.(23)

RESULTS

Of 26,914 patients in the first version of the registry, 8,029 met inclusion criteria for this analysis (≥ 18 years old, SARS-CoV-2+, hospitalized from the ED). We excluded 59 who were missing vital status at 30 days and 9 who died in the ED prior to vital signs, leaving 7,961 in the analysis cohort. Of the 7,961, 1847 (23.2%) died within 30 days. The median age of patients in the cohort was 66 years (interquartile range 54-77); 55.6% were male; and 58.7% had at least one comorbid condition. Of note, the median oxygen saturation was 92% (IQR 87% to 95%) overall and 87% (IQR 77% to 94%) in those who died ($p < 0.0001$).

Table 2. Patient characteristics, clinical characteristics, and 30-day mortality of SARS-CoV-2+ patients hospitalized from the Emergency Department.

	Total	Median (IQR)	Deaths	Median (IQR)
Overall	**7,961**		**1,847**	
Key clinical measures				
Age, in years	7,961	66 (54 - 77)	1,847	75 (65 - 84)
Oxygen saturation, room air (%)	7,961	92 (87 - 95)	1,847	87 (77 - 94)
C-reactive protein (mg/dL)	4,809	10.9 (5.1 - 19.5)	1,043	16.9 (9.3 - 25.1)
Body Mass Index*	7,059	28.4 (24.5 - 33.7)	1,580	28 (24 - 33)
Other vital signs				
Temperature (°C)†	7,960	37.2 (36.7 - 37.9)	1,847	37.1 (36.7 - 37.9)
Heart rate (beats/minute)‡	7,956	96 (83 - 110)	1,843	96 (82 - 111)
Systolic blood pressure (mm Hg)	7,961	130 (116 - 146)	1,847	128 (112 - 147)
Respiratory rate (breaths/minute)	7,954	20 (18 - 24)	1,845	20 (18 - 26)
Other blood tests				
White blood cell count (10^3/uL)	7,909	7.1 (5.3 - 10)	1,829	8.4 (6 - 12)
Lymphocyte count (10^3 cells/uL)	3,437	1.1 (0.7 - 1.8)	620	1 (0.6 - 1.5)
Hemoglobin (g/dL)	7,911	13 (11.9 - 14.5)	1,832	13 (11 - 14.4)
Platelets (10^3 cells/uL)	7,901	206 (158 - 270)	1,829	196 (147 - 259)
Sodium (mEq/L)	7,452	136 (133 - 139)	1,712	137 (133 - 142)
Potassium (mEq/L)	7,878	4.2 (3.8 - 4.6)	1,827	4.4 (3.9 - 4.9)
BUN (mg/dL)	7,823	19 (12 - 33)	1,810	30 (19 - 51)
Creatinine (mg/dL)	3,619	1.1 (0.8 - 1.5)	664	1.4 (1 - 2.1)
Total bilirubin (mg/dL)	7,111	0.6 (0.4 - 0.8)	1,660	0.6 (0.4 - 0.9)

IQR=Interquartile range
* P-value for rank sum test comparison of died versus survived; BMI, p=0.0008
† Temperature, p=0.009
‡ Heart rate, p=0.750
All other p-values < 0.0001

Of the demographic risk factors, age group, male sex, and residence in a nursing home/assisted living facility were the principal mortality predictors (Table 2). In the multivariable model, age 80+ years increased 30-day mortality risk by a factor of 4.92 (95% CI 3.81 - 6.03); male sex increased it by 1.11 (95% CI 1.03 - 1.20); and nursing home/assisted living facility residence increased it by 1.37 (95% CI 1.24 - 1.51). On a univariable basis, Hispanic ethnicity and smoking status were associated with lower mortality risk, but after adjusting for other variables, including the younger age of Hispanics and smokers, the risk ratio for mortality for Hispanic ethnicity was 0.96 (95% CI 0.85 - 1.06) and for smoking was 0.88 (95% CI 0.74 - 1.02).

In the univariable analysis, extreme obesity (BMI ≥ 40) did not increase risk, but after adjustment for age, sex, and other comorbidities, the risk ratio for BMI ≥ 40 was 1.30 (95% CI 1.15-1.44). In addition to obesity, the multivariable analysis (Table 3) showed that other comorbidities associated with increased risk of death were chronic cardiac disease, chronic pulmonary disease, liver disease as indicated by a total bilirubin ≥ 2.0 mg/dL, and kidney disease as indicated by a creatinine ≥ 1.2 mg/dL or BUN ≥ 40. Asthma and diabetes were not significant risk factors. Patients who arrived from a nursing home had an increased risk of mortality (Risk Ratio 1.37, 95% CI 1.24 - 1.51).

Table 3. Effect of patient and clinical characteristics on 30-day mortality in SARS-CoV-2+ patients hospitalized from the Emergency Department.

		Total	% of non-missing	Deaths	30-day Mortality	Risk Ratio	Multivariable Risk Ratio (95% CI)
Overall		**7,961**		**1,847**	**23.2%**		
Age Group (years)	**18 - 49**	1,462	18.4%	82	5.6%	ref	ref
	50 - 59	1,365	17.2%	171	12.5%	2.23	1.87 (1.42 - 2.31)
	60 - 69	1,918	24.1%	417	21.7%	3.88	2.92 (2.28 - 3.57)
	70 - 79	1,631	20.5%	491	30.1%	5.37	3.93 (3.06 - 4.79)
	80+	1,585	19.9%	686	43.3%	7.72	4.92 (3.81 - 6.03)
Sex at birth	**Female**	3,533	44.4%	784	22.2%	ref	ref
	Male	4,428	55.6%	1,063	24.0%	1.08	1.11 (1.03 - 1.2)
Race/ethnicity	**White (non H/L)**	2,130	26.8%	570	26.8%	ref	ref
	Asian (non H/L)	267	3.4%	66	24.7%	0.92	1.05 (0.84 - 1.27)
	Black (non H/L)	2,630	33.0%	590	22.4%	0.84	1.06 (0.96 - 1.16)
	Hispanic/Latinx (H/L)	1,946	24.4%	349	17.9%	0.67	0.96 (0.85 - 1.06)
	Other or Unknown	988	12.4%	272	27.5%	1.03	1.18 (1.05 - 1.32)
Resides in Nursing Home or Assisted Living		1,110	13.9%	478	43.1%	2.16	1.37 (1.24 - 1.51)
Smoker		528	7.5%	87	16.5%	0.74	0.88 (0.74 - 1.02)
***Body Mass Index (BMI)**	**18.5 - <40**	6,084	86.2%	1,348	22.2%	ref	ref
	< 18.5	225	3.2%	67	29.8%	1.34	1.05 (0.87 - 1.22)
	≥ 40	750	10.6%	165	22.0%	0.99	1.3 (1.15 - 1.44)
Comorbidities, among 7 with *	**0**	3,285	41.3%	640	19.5%	ref	
	1	2,490	31.3%	566	22.7%	1.17	
	2 +	2,186	27.5%	641	29.3%	1.51	
Asthma		673	8.5%	129	19.2%	0.81	0.99 (0.85 - 1.13)
***Cancer**		736	9.3%	202	27.4%	1.20	1 (0.89 - 1.12)
***Chronic Cardiac Disease**		1,576	19.8%	514	32.6%	1.56	1.13 (1.03 - 1.23)
	Atrial Fibrillation	746	9.4%	261	35.0%	1.59	
	Heart Disease	545	6.9%	157	28.8%	1.26	

Heart Failure		805	10.1%	257	31.9%	1.44	
***Chronic Pulmonary Disease**		718	9.0%	232	32.3%	1.45	1.16 (1.03 - 1.29)
	COPD	585	7.4%	190	32.5%	1.45	
	Bronchiectasis	22	0.5%	7	31.8%	1.16	
	Other Lung Disease	176	2.2%	54	30.7%	1.33	
	Pulmonary Fibrosis	28	0.6%	11	39.3%	1.43	
***Diabetes**		2,528	31.8%	635	25.1%	1.12	1.04 (0.96 - 1.12)
***Liver Disease (Total Bilirubin ≥2.0)**		211	3.0%	77	36.5%	1.59	1.48 (1.25 - 1.7)
***Kidney Disease (Creatinine ≥1.2)**		1,459	40.3%	423	29.0%	2.60	
High Creatinine (≥1.2) or BUN (≥40)		2,457	31.3%	882	35.9%	2.08	1.46 (1.34 - 1.57)
Respiratory rate (breaths/min)	**< 20**	3,292	41.4%	589	17.9%	ref	ref
	20 - 29	3,872	48.7%	931	24.0%	1.34	1.18 (1.08 - 1.28)
	≥ 30	790	9.9%	325	41.1%	2.30	1.71 (1.52 - 1.89)
Oxygen saturation, room air (%)	**≥ 92**	4,627	58.1%	669	14.5%	ref	ref
	< 92	3,334	41.9%	1,178	35.3%	2.44	1.91 (1.75 - 2.06)
Altered mental status		1,268	15.9%	342	27.0%	1.20	
BUN (mg/dL)	**< 20**	4,058	51.9%	483	11.9%	ref	
	20 - 39	2,231	28.5%	661	29.6%	2.49	
	≥ 40	1,534	19.6%	666	43.4%	3.65	
C-reactive protein (mg/dL)	**< 5.0**	1,161	24.1%	104	9.0%	ref	ref
	5.0 - 9.9	1,090	22.7%	185	17.0%	1.89	1.15 (0.99 - 1.31)
	≥ 10.0	2,558	53.2%	754	29.5%	3.29	1.59 (1.41 - 1.77)

ref=reference; CI=Confidence Interval; COPD=Chronic Obstructive Pulmonary Disease

Missing over 1.75%: BMI=11.3%; Smoker=11.7%; Bronchiectasis=45.2%; Pulmonary Fibrosis=45.2%; Bilirubin=10.7%; Creatinine=54.5%; C-reactive protein=39.6%

Relative risk is risk of death relative to reference if indicated, otherwise to not having risk factor.

Table 3 also shows that increase in respiratory rate, decrease in oxygen saturation, and increase in CRP each corresponded with stepwise increase in risk of mortality.

Compared with the 4C validation dataset from Knight et al, the mean 4C Mortality Scores were lower in our dataset (mean score 9.7 vs. 10.6). (Figure 1A). Using nine 4C score categories, the AUROC from the RECOVER dataset was comparable to that of the original 4C validation dataset. Using nine 4C score categories, the AUROC from the RECOVER dataset was not different than the original 4C validation dataset (AUROC: RECOVER 0.779 (95% CI 0.768 - 0.790) vs 4C validation 0.763 [95% CI 0.757 - 0.769). (Figure 1B). Our observed category-specific mortalities were lower than those in the 4C validation dataset. Using the mortalities from the 4C validation dataset would have over-estimated risk by

2.5% on average. (Mean prediction error 2.5%. √Calibration Error 0.029, and √Brier Score 0.383). (Table 4). (Figure 1C).

Table 4. Comparison of observed mortality by 4C mortality risk group for RECOVER Dataset of SARS-CoV-2+ patients hospitalized from the Emergency Department.

4C Mortality Risk Group	With CRP*			Without CRP		
	Predicted Risk	Observed Mortality	Prediction Error	Predicted Risk	Observed Mortality	Prediction Error
Overall	**25.7%**	**23.2%**	**2.5%**	**19.9%**	**23.2%**	**-3.3%**
Low (0-3 points)	1.3%	1.1%	0.2%	1.2%	2.0%	-0.8%
Intermediate (4-8 points)	9.8%	7.0%	2.8%	9.7%	11.2%	-1.5%
High (9-14 points)	30.4%	27.9%	2.5%	29.3%	34.5%	-5.3%
Very high (≥15 points)	60.2%	57.6%	2.9%	57.5%	62.9%	-5.7%

	AUROC	√Calibration Error	√Brier Score
4C Validation	0.763 (0.757 - 0.769)		
RECOVER with CRP	0.779 (0.768 - 0.790)	0.0290	0.3830
RECOVER without CRP	0.769 (0.758 - 0.780)	0.0421	0.3874

* CRP=C-reactive protein.

Dropping CRP from the 4C Mortality Score reduced the scores overall (mean score 8.4) but did not substantially change discrimination (AUROC 0.769, 95% CI 0.758 - 0.780). Dropping the CRP component did affect calibration. The category-specific mortalities in our dataset were now higher than those in the 4C validation dataset. Using the mortalities from the 4C validation dataset would have under-estimated risk by 3.3% on average. (Mean prediction error -3.3%. √Calibration Error 0.042, and √Brier Score 0.387).

DISCUSSION

In this analysis of multicenter data from the RECOVER Network, our results confirmed several previous findings for risk factors for COVID-19 mortality, including older age, comorbidities, BMI ≥ 40 kg/m^2, higher respiratory rate, and lower oxygen saturation. (4-7, 9, 11-14, 24) In addition, as reported by Graselli et al in critically ill patients, we observed that male sex is predictive of mortality.(7). We also observed the expected, but previously unquantified finding that arrival to the emergency care setting from a nursing home was associated with increased mortality. While this has not been specifically mentioned in

other studies, Ferrando-Vivas et al found that functional dependence was related to mortality (Hazard Ratio 1.425).(4)

In the RECOVER Network, COVID-19 related hospitalizations are higher among SARS-Cov-2+ Hispanic patients when compared to Non-Hispanics, but the adjusted mortality is similar to non-Hispanic whites. (25) Similarly, Mackey et al reported that hospitalizations for COVID-19 among those who identify their ethnicity as Hispanic were proportionately higher than for their non-Hispanic white counterparts(26) but the case fatality rate was similar between Hispanic and non-Hispanic patients.

We also found that the comorbid conditions such as liver disease defined as elevated total bilirubin ≥2.0 and kidney disease defined as creatinine ≥1.2 mg/dl or BUN ≥40 had an independent association with 30-day mortality in hospitalized SARS-Cov-2+ patients. Surprisingly, previous studies and our results did not establish diabetes as a significant risk factor. (26-28) But our findings on the association of smoking with 30-day mortality did not concur with previous studies. Smoking as well as cumulative smoking exposure was predictive of mortality in previous studies (29) but, we did not find a statistically significant association after controlling for other variables. Finally, among the clinical variables, tachypnea (respiratory rate ≥20) and hypoxemia (oxygen saturation <92%) were significant predictors of mortality. Zhou et al reported higher odds of mortality (adjusted OR 4.x) for an oxygen saturation <92%. (13)

Given the multiplicity of variables associated with 30-day mortality, clinicians need a simple score to better predict short-term mortality. The 4C Mortality Score is one such score and it performed well in our dataset. Discrimination was excellent, and calibration was also good, although using the category-specific mortalities from the 4C validation dataset would have over-estimated risk. CRP was missing in 40% of the records in our study, so we examined the performance of the 4C Mortality Score without the CRP component. Although discrimination remained good, the category-specific risks from the 4C validation were now too low. When CRP was removed from the score, many patients with high CRP values moved into a lower risk category. Those patients who remained with high 4C Mortality Scores despite removal of CRP died at a higher rate than those whose risk score decreased, but those with high CRP values who moved to a lower risk group had higher mortality than the average for their new lower risk group. This might be referred to as stage migration effect. When the high CRP patients moved from the very-high risk group to the high-risk group, the average mortality went up in both groups. Based on our observations, we suggest using the 4C Mortality Score without the CRP component, but recalibrating risk estimates as per our Table 4 or Supplementary Table A. Using category-specific risks as opposed to the 4 risk groups (low, intermediate, high, very high) is preferred because it doesn't assume the distribution

across the risk groups is the same in different populations. This modified 4C Mortality Score could assist with triage decisions, to inform patients and their family members of prognostic information, and to help with forecasting of resource utilization in the hospital.

Limitations

This is a national study of hospitalized SARS-Cov-2+ patients. The large sample size, the number and diversity of the participating sites, and a comprehensive list of data elements are major strengths of this study. However, some sites contributed more SARS-Cov-2+ patients than others. We did see regional differences in 30-day mortality, but these did not affect the risk ratios. As noted above, CRP was missing in 40% of patients.

CONCLUSIONS

We conclude that among SARS-Cov-2+ hospitalized patients, older patients with comorbid conditions and those with hypoxemia at the time of presentation have a very high risk of dying within 30 days. We independently validate the 4C Mortality Score as predicting risk of death in hospitalized SARS-Cov-2+ patients, but we recommend dropping the CRP component of the score and using our recalibrated mortality risk estimates.

Conflict of Interests: None

Grants: Funded by Department of Emergency Medicine, Indiana University School of Medicine. No Grant Number.

Author contributions: PG, MAK, AJG, CB, CC and JK conceived and designed the study, JK obtained research funding, designed, and organized the registry. MAK, LM provided statistical advice on study design and analyzed the data; AJG, MAK, PG drafted the manuscript, and all authors contributed substantially to its revision. PG takes responsibility for the paper as a whole.

Data sharing agreement: No data is available

Ethics approval statement: The RECOVER registry protocol was reviewed by the institutional review boards (IRBs) at all sites; 3 IRBs provided approval with waiver of informed consent. All others provided an exemption from human subjects designation. All data were anonymized prior to analysis. Local IRB: 56234

REFERENCES:

1. Lipsitch M, Swerdlow DL, Finelli L. Defining the Epidemiology of Covid-19 — Studies Needed. New England Journal of Medicine. 2020;382(13):1194-6.
2. Kerlin MP, Costa DK, Davis BS, Admon AJ, Vranas KC, Kahn JM. Actions Taken by US Hospitals to Prepare for Increased Demand for Intensive Care During the First Wave of COVID-19: A National Survey. LID - S0012-3692(21)00478-5 [pii] LID - 10.1016/j.chest.2021.03.005 [doi]. (1931-3543 (Electronic)).
3. Yang X, Yu Y, Xu J, Shu H, Xia J, Liu H, et al. Clinical course and outcomes of critically ill patients with SARS-CoV-2 pneumonia in Wuhan, China: a single-centered, retrospective, observational study. (2213-2619 (Electronic)).
4. Ferrando C, Mellado-Artigas R, Gea A, Arruti E, Aldecoa C, Bordell A, et al. Patient characteristics, clinical course and factors associated to ICU mortality in critically ill patients infected with SARS-CoV-2 in Spain: A prospective, cohort, multicentre study. Rev Esp Anestesiol Reanim. 2020;67(8):425-37.
5. Ferrando-Vivas P, Doidge J, Thomas K, Gould DW, Mouncey P, Shankar-Hari M, et al. Prognostic Factors for 30-Day Mortality in Critically Ill Patients With Coronavirus Disease 2019: An Observational Cohort Study. Crit Care Med. 2021;49(1):102-11.
6. Gallo Marin B, Aghagoli G, Lavine K, Yang L, Siff EJ, Chiang SS, et al. Predictors of COVID-19 severity: A literature review. Reviews in Medical Virology. 2021;31(1):e2146.
7. Grasselli G, Greco M, Zanella A, Albano G, Antonelli M, Bellani G, et al. Risk Factors Associated With Mortality Among Patients With COVID-19 in Intensive Care Units in Lombardy, Italy. JAMA Internal Medicine. 2020;180(10):1345-55.
8. Liang W, Liang H, Ou L, Chen B, Chen A, Li C, et al. Development and Validation of a Clinical Risk Score to Predict the Occurrence of Critical Illness in Hospitalized Patients With COVID-19. JAMA Internal Medicine. 2020;180(8):1081-9.
9. Ruan Q, Yang K, Wang W, Jiang L, Song J. Clinical predictors of mortality due to COVID-19 based on an analysis of data of 150 patients from Wuhan, China. Intensive Care Med. 2020;46(5):846-8.
10. Wynants L, Van Calster B, Collins GS, Riley RD, Heinze G, Schuit E, et al. Prediction models for diagnosis and prognosis of covid-19: systematic review and critical appraisal. BMJ. 2020;369:m1328.

11. Xu J, Yang X, Yang L, Zou X, Wang Y, Wu Y, et al. Clinical course and predictors of 60-day mortality in 239 critically ill patients with COVID-19: a multicenter retrospective study from Wuhan, China. Crit Care. 2020;24(1):394.
12. Zhang S, Guo M, Duan L, Wu F, Hu G, Wang Z, et al. Development and validation of a risk factor-based system to predict short-term survival in adult hospitalized patients with COVID-19: a multicenter, retrospective, cohort study. Crit Care. 2020;24(1):438.
13. Zhao Z, Chen A, Hou W, Graham JM, Li H, Richman PS, et al. Prediction model and risk scores of ICU admission and mortality in COVID-19. PLoS One. 2020;15(7):e0236618.
14. Zhou F, Yu T, Du R, Fan G, Liu Y, Liu Z, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. The Lancet. 2020;395(10229):1054-62.
15. Levy TJ, Richardson S, Coppa K, Barnaby DP, McGinn T, Becker LB, et al. A predictive model to estimate survival of hospitalized COVID-19 patients from admission data. medRxiv. 2020:2020.04.22.20075416.
16. Brajer N, Cozzi B, Gao M, Nichols M, Revoir M, Balu S, et al. Prospective and External Evaluation of a Machine Learning Model to Predict In-Hospital Mortality of Adults at Time of Admission. JAMA Network Open. 2020;3(2):e1920733-e.
17. Knight SR, Ho A, Pius R, Buchan I, Carson G, Drake TM, et al. Risk stratification of patients admitted to hospital with covid-19 using the ISARIC WHO Clinical Characterisation Protocol: development and validation of the 4C Mortality Score. BMJ. 2020;370:m3339.
18. Kline JA, Pettit KL, Kabrhel C, Courtney DM, Nordenholz KE, Camargo CA. Multicenter registry of United States emergency department patients tested for SARS-CoV-2. J Am Coll Emerg Physicians Open. 2020;1(6):1341-8.
19. Niknam K KM, Mitarai T, Nesbitt J, Wilson JG, editor Emergency Critical Care SOFA Score Predicts In-Hospital Mortality. Society of Academic Emergency Medicine; 2019; Las Vegas.
20. StataCorp. Stata Multiple-Imputation Reference Manual College Station, TX: Stata Press; 2021 [Release 17:[Available from: https://www.stata.com/manuals/mi.pdf.
21. Cummings P. Estimating Adjusted Risk Ratios for Matched and Unmatched Data: An Update. The Stata Journal. 2011;11(2):290-8.
22. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics. 1988;44(3):837-45.
23. Newman TB, Kohn MA. Evidence-based diagnosis : an introduction to clinical epidemiology. Cambridge ; New York: Cambridge University Press,; 2019.
24. Liang W, Liang H, Ou L, Chen B, Chen A, Li C, et al. Development and Validation of a Clinical Risk Score to Predict the Occurrence of Critical Illness in Hospitalized Patients With COVID-19. JAMA Internal Medicine. 2020;180(8):1081-9.
25. Kline JA, Camargo CA, Jr., Courtney DM, Kabrhel C, Nordenholz KE, Aufderheide T, et al. Clinical prediction rule for SARS-CoV-2 infection from 116 U.S. emergency departments 2-22-2021. PLOS ONE. 2021;16(3):e0248438.
26. Mackey K, Ayers CK, Kondo KK, Saha S, Advani SM, Young S, et al. Racial and Ethnic Disparities in COVID-19-Related Infections, Hospitalizations, and Deaths : A Systematic Review. Ann Intern Med. 2021;174(3):362-73.
27. Mehta PA-O, Rushakoff RJ. A Diabetes and COVID-19 Dilemma: Balancing Parents' Health & Children's Education. (1932-2968 (Electronic)).
28. Mehta PA-O, Kohn MA, Koliwad S, Rushakoff RJ. Lack of association between either outpatient or inpatient glycemic control and COVID-19 illness severity or mortality in patients with diabetes. . BMJ Open. 2021.
29. Lowe KE, Zein J, Hatipoğlu U, Attaway A. Association of Smoking and Cumulative Pack-Year Exposure With COVID-19 Outcomes in the Cleveland Clinic COVID-19 Registry. JAMA Internal Medicine. 2021.

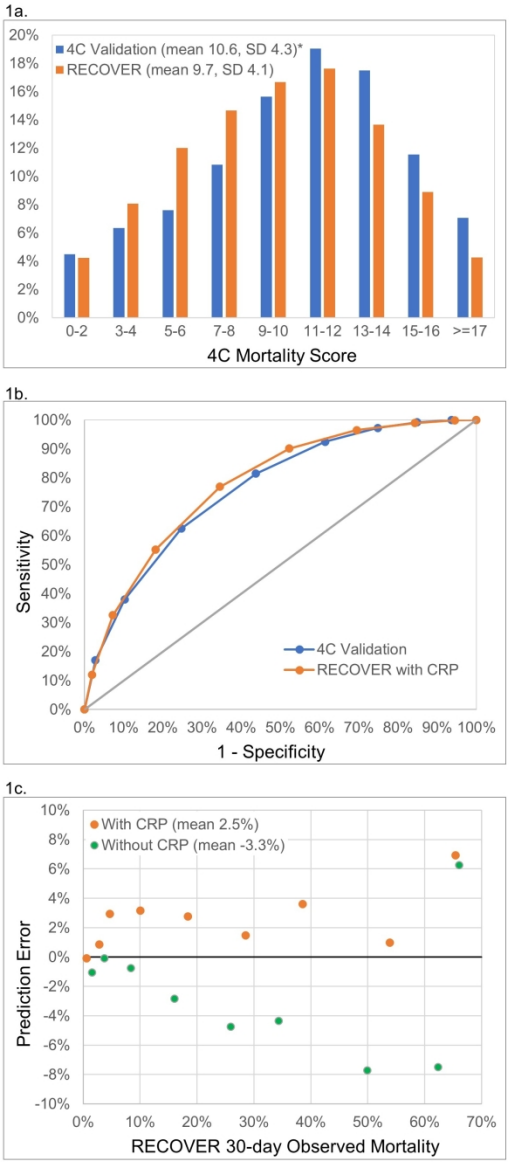

Figure 1: 1a. 4C Mortality Scores were lower in the RECOVER dataset than in the original 4C validation dataset. *The mean and SD in the validation dataset were estimated from the original 4C report. 1b. ROC curves for the 4C Mortality Score (categorized into the 9 ranges from Figure 1a) in the 4C validation dataset and the RECOVER dataset. AUROC: RECOVER 0.779 (95% CI 0.768 - 0.790), 4C validation 0.763 [95% CI 0.757 - 0.769). 1c. Calibration plot (modified Bland-Altman) showing prediction error versus observed mortality for the 4C Mortality Score with and without the C-reactive protein (CRP) component.

215x279mm (300 x 300 DPI)

Table A. Observed mortality by 4C Risk Group calculated without C-reactive protein for RECOVER dataset of SARS-CoV-2+ patients hospitalized from the Emergency Department.

4C Mortality Risk Group	Patient Distribution	Predicted Risk	Observed Mortality	Prediction Error
0-2	8.1%	0.5%	1.6%	-1.1%
3-4	10.8%	3.7%	3.7%	-0.1%
5-6	14.4%	7.6%	8.4%	-0.7%
7-8	16.9%	13.2%	16.0%	-2.8%
9-10	17.9%	21.2%	25.9%	-4.7%
11-12	16.3%	30.0%	34.4%	-4.4%
13-14	10.3%	42.2%	49.9%	-7.7%
15-16	4.7%	54.9%	62.4%	-7.5%
≥17	0.7%	72.3%	66.1%	6.3%

Mean Prediction Error	-3.3%
√Calibration Error	0.0421
√Brier Score	0.3875

STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No	Recommendation
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract "retrospective cohort study" √ **(page 3)**
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found√ **(page 3)**
Introduction		
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported√**(page 6)**
Objectives	3	State specific objectives, including any prespecified hypotheses√**(page6,7)**
Methods		
Study design	4	Present key elements of study design early in the paper√ **(page6)**
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection **(page 3,7)**
Participants	6	**(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up √(page 6)** *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed *Case-control study*—For matched studies, give matching criteria and the number of controls per case
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable √ **(page 7)**
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group √ **(page 7)**
Bias	9	Describe any efforts to address potential sources of bias √ **(page 3, 12)**
Study size	10	Explain how the study size was arrived at √ **(page 7)**
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why √ **(page 7-9)**
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding √ **(page 8)**
		(*b*) Describe any methods used to examine subgroups and interactions √ **(page 8)**
		(*c*) Explain how missing data were addressed √ **(page 8)**
		(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed *Case-control study*—If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy
		(*e*) Describe any sensitivity analyses

Continued on next page

Results		
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed √
		(b) Give reasons for non-participation at each stage
		(c) Consider use of a flow diagram
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders √
		(b) Indicate number of participants with missing data for each variable of interest
		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount) 30 day mortality √
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time **√ (page 7)**
		Case-control study—Report numbers in each exposure category, or summary measures of exposure
		Cross-sectional study—Report numbers of outcome events or summary measures
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included **√ (page 9,11)**
		(*b*) Report category boundaries when continuous variables were categorized
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses **√ (page 9,10)**
Discussion		
Key results	18	Summarise key results with reference to study objectives**√ (page 9)**
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias**√(page 12)**
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence**√ (page 12)**
Generalisability	21	Discuss the generalisability (external validity) of the study results**√ (page 12)**
Other information		
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based**√ (page 2)**

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

BMJ Open

External Validation of the 4C Mortality Score for Hospitalized COVID-19 Patients in the RECOVER network

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2021-054700.R1
Article Type:	Original research
Date Submitted by the Author:	24-Nov-2021
Complete List of Authors:	Gordon, Alexandra June; Stanford University School of Medicine, Emergency Medicine Govindarajan, Prasanthi; Stanford University School of Medicine, Department of Emergency Medicine Bennett, Christopher L.; Stanford University School of Medicine, Department of Emergency Medicine Matheson, Loretta ; Stanford University School of Medicine, Department of Emergency Medicine Kohn, Michael; Stanford University School of Medicine, Department of Emergency Medicine Camargo, Carlos; Massachusetts General Hospital, Harvard Medical School, Emergency Medicine Kline, Jeffrey ; Indiana University School of Medicine, Emergency
Primary Subject Heading:	Emergency medicine
Secondary Subject Heading:	Infectious diseases
Keywords:	COVID-19, ACCIDENT & EMERGENCY MEDICINE, Adult intensive & critical care < ANAESTHETICS, EPIDEMIOLOGY, GENERAL MEDICINE (see Internal Medicine)

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

External Validation of the 4C Mortality Score for Hospitalized COVID-19 Patients in the RECOVER network

Alexandra June Gordon, MD (First author)
Department of Emergency Medicine, Stanford University School of Medicine,
ajgordon@stanford.edu

Prasanthi Govindarajan, MBBS, MAS (Co-first author)
Department of Emergency Medicine, Stanford University School of Medicine
prashagovindarajan@stanford.edu

Christopher L. Bennett, MD
Department of Emergency Medicine, Stanford University School of Medicine
cleebennett@stanford.edu

Loretta Matheson, MS
Department of Emergency Medicine, Stanford University School of Medicine
lorettamatheson@stanford.edu

Michael A. Kohn, MD
Department of Emergency Medicine, Stanford University School of Medicine
Michael.Kohn@ucsf.edu

Carlos A. Camargo, Jr., MD, DrPH
Department of Emergency Medicine, Massachusetts General Hospital, Harvard Medical School, Boston, MA
CCAMARGO@PARTNERS.ORG

Jeffrey A. Kline MD
Department of Emergency Medicine, Indiana University School of Medicine
jkline@wayne.edu

Corresponding author: Jeffrey Kline MD, jkline@wayne.edu
Department of Emergency Medicine,
Indiana University School of Medicine
5th/3rd Bank Building, 3rd Floor
640 Eskenazi Avenue
Indianapolis, IN 46202
Telephone: 317-670-0541

Manuscript word count (not including title, abstract, acknowledgment, references, tables, and figure legends): 2707

ABSTRACT:

Objectives: Estimating mortality risk in hospitalized SARS-CoV-2+ patients may help with choosing level of care and discussions with patients. The Coronavirus Clinical Characterisation Consortium Mortality Score (4C Score) is the most promising COVID-19 mortality risk model. We examined the association of risk factors with 30-day mortality in hospitalized, full-code SARS-CoV-2+ patients and investigated the discrimination and calibration of the 4C Score. This was a retrospective cohort study of SARS-CoV-2+ hospitalized patients within the RECOVER (REgistry of suspected COVID-19 in EmeRgency care) Network.

Setting: 99 emergency departments (EDs) across the US.

Participants: Patients ≥ 18 years old, positive for SARS-CoV-2 in the ED, and hospitalized.

Primary outcome: Death within 30 days of the index visit. We performed logistic regression analysis, reporting multivariable risk ratios (MVRRs) and calculated the AUROC and mean prediction error for the original 4C Score and after dropping the c-reactive protein component (CRP).

Results: Of 6,802 hospitalized COVID-19 patients, 1,149 (16.9%) died within 30 days. The 30-day mortality was increased with age 80+ years (MVRR = 5.79, 95% CI 4.23 - 7.34); male sex (MVRR = 1.17,1.05 - 1.28); and nursing home/assisted living facility residence (MVRR = 1.29, 1.1 - 1.48). The 4C Score had comparable discrimination in the RECOVER dataset compared with the original 4C validation dataset (AUROC: RECOVER 0.786 (95% CI 0.773 - 0.799), 4C validation 0.763 (95% CI 0.757 - 0.769). Score-specific mortalities in our sample were lower than in the 4C validation sample (mean prediction error 5.97%). Dropping the CRP component from the 4C Score did not substantially affect discrimination and 4C risk estimates were now close (mean prediction error 0.7%).

Conclusions: We independently validated 4C Score as predicting risk of 30-day mortality in hospitalized SARS-CoV-2+ patients. We recommend dropping the CRP component of the score and using our recalibrated mortality risk estimates.

ARTICLE SUMMARY:

Strengths and Limitations:

- In this first study using a national US sample of patients who tested positive for SARS-CoV-2 and were hospitalized through emergency departments, our results confirmed the previous findings that older age, comorbidities, BMI ≥ 40 kg/m2, higher respiratory rate, and lower oxygen saturation were associated with 30-day mortality.

-We also observed that the arrival to the emergency care setting from a nursing home was associated with increased mortality.

-We independently validated 4C Mortality Score as predicting risk of 30-day mortality in hospitalized SARS-CoV-2+ patients.

-We recommend dropping the CRP component of the score and using our recalibrated mortality risk estimates when estimating the 30-day mortality in hospitalized patients who test positive for SARS-CoV-2.

INTRODUCTION

The COVID-19 pandemic has placed tremendous strain on emergency and critical care resources in hospitals worldwide.[1–3] To prepare the healthcare systems for the surges, several studies have developed prediction models to assess mortality risk in patients hospitalized with COVID-19. These studies identified the following risk factors for mortality or critical care admission: age, sex, comorbid conditions, laboratory values, and vital signs.[4–16]

In a systematic review that evaluated many of these risk prediction studies using the prediction model risk of bias assessment tool (PROBAST), Wynants et al. concluded that many of the current risk models may be misleading.[10] However, the authors' analysis suggested that one COVID-19 mortality prediction model, the 4C Mortality Score, which was built on a large UK data set, had relatively low risk of bias in most domains by the PROBAST criteria. The 4C Mortality Score includes eight variables - age, sex, respiratory rate, oxygen saturation, number of comorbidities, level of consciousness, blood urea nitrogen, and C-reactive protein (CRP); see table 1.[17] While there has been continued interest in the development of prediction models for COVID-19, the 4C Mortality Score represented one of the first with a low risk of bias and therefore a good candidate for verification in other populations.

Table 1. Point assignment for 4C Mortality Score

Age Group (years)	18 - 49	
	50 - 59	+ 2
	60 - 69	+ 4
	70 - 79	+ 6
	80+	+ 7
Sex at birth	Female	
	Male	+ 1
Comorbidities*	0	
	1	+ 1
	2 +	+ 2
Respiratory rate (breaths/minute)	< 20	
	20 - 29	+ 1
	≥ 30	+ 2
Oxygen saturation, room air (%)	≥ 92	
	< 92	+ 2
Altered mental status †	No	
	Yes	+ 2

BUN (mg/dL)§	< 20	
	20 - 39	+ 1
	≥ 40	+ 3
C-reactive protein (mg/dL)§	< 5.0	
	5.0 - 9.9	+ 1
	≥ 10.0	+ 2

* Comorbidities: High Body Mass Index, Cancer, Chronic cardiac disease, Chronic pulmonary disease, Diabetes, Liver disease, Kidney disease.

† Altered mental status is patient-reported symptom whereas 4C used Glasgow coma scale < 15.

§ BUN and C-reactive protein units converted from 4C.

In this study, we investigated the risk of 30-day mortality in hospitalized SARS-CoV-2+ patients within the RECOVER (REgistry of suspected COVID-19 in EmeRgency care) Network.[18] In a large cohort of SARS-CoV-2+ patients hospitalized from 99 US emergency departments (EDs), we determined the relation of demographic and clinical factors with 30-day mortality and investigated the discrimination and calibration of the 4C Mortality Score with and without the CRP value.

METHODS

In this retrospective cohort study, we included patient-level data from the RECOVER Network, a national registry of patients who were tested for SARS-CoV-2 during their ED visit. We restricted the analysis to full code status patients ≥ 18 years old who tested positive for SARS-CoV-2 and were hospitalized from the ED.[18] The study was approved or deemed exempted by the Institutional Review Boards of all participating sites.

Data Source

We obtained data from 40 medical centers representing 99 EDs from 27 US states and the District of Columbia. Data were collected using a REDCap data collection form that was distributed to the ED sites during the study period (3/2020-9/2020); our data were downloaded from the registry in 12/2020. The REDCap form (Appendix A) had seven sections and 204 questions, which generated 360 data elements. Variables reflected a combination of routinely collected information (e.g. patient demographics, medical history, vital signs, and diagnostic test results), patient-reported symptoms and risk exposures, clinical outcomes (e.g. admission, therapies, death), and those deemed important by the RECOVER Network

steering committee. After creation, but prior to launch, the data form was piloted at 19 sites and refined. For additional section details and the questions, please refer to the data collection form in the appendix (supplement). The data were obtained from the electronic healthcare record using a combination of electronic download for routinely collected, coded variables (e.g., age, vital signs and laboratory values), supplemented by chart review by research personnel, using methods previously described.[18]

Patient Involvement

Patients were not involved in the development of our work, setting the research question or determining the outcome measures. This applies to both the RECOVER Network and the work presented here. Given the nature and limitations of emergency care during the COVID-19 pandemic, it was not appropriate or possible to involve patients or the public in the design, or conduct, reporting, or dissemination plans of our work.

Study Variables

We analyzed patient characteristics such as demographics, vital signs, symptoms, risks for infection, comorbidities, and medications. Following the 4C Mortality Score, we categorized the patients into five age groups (18-49, 50-59, 60-69, 70-79, 80+). The US standard ethnicity (Latinx/Hispanic yes/no) and race categories were combined into 8 categories (Non-Hispanic White, Non-Hispanic African American/Black, Non-Hispanic Asian, Non-Hispanic Native American/Alaskan Native, Non-Hispanic Pacific Islander/Native Hawaiian, Non-Hispanic mixed, and Non-Hispanic Unknown). In the analysis, we combined Non-Hispanic Native American (0.2%), Non-Hispanic Pacific Islander (0.2%), Non-Hispanic mixed, Other, and Unknown into a single group (12.8%).

All included patients had a positive reverse transcriptase polymerase chain reaction test (RT-PCR) test for SARS-CoV-2. Almost all the tests were performed during the ED visit, but we also included patients who had a test in a physician's office or urgent care immediately prior to the ED visit. We excluded patients whose 30-day vital status could not be ascertained, those who died in the ED before vital signs were recorded, and those who did not have full code status.

Study Outcome

The primary outcome was death within 30 days of the index visit. The 4C mortality score's predictive accuracy was measured by the area under the ROC curve (AUROC) and mean prediction error.

Data Preparation and Statistical Analysis

For comparison with the other cohorts, we report the median and interquartile range of continuous variables – both in the entire cohort and in the subgroup who died within 30 days – and compared the median values using the rank sum test. We performed univariable analysis on 26 independent variables that were included on the data collection form using complete (non-missing) data and reporting risk ratios for 30-day mortality. For risk ratio reporting of continuous variables, we chose category boundaries based on the 4C Mortality Score (age, respiratory rate, oxygen saturation, BUN, and CRP) or other published mortality prediction models (BMI, creatinine, total bilirubin).[19]

We selected variables for our multivariable logistic regression model based on the 4C Mortality Score, other prior studies, and clinical judgement. The RECOVER dataset has complete data (< 1.5% missing) on most variables, with the exception of CRP, body mass index (BMI), bilirubin, and smoking status. For our multivariable analysis, we used imputed values for missing data using Stata's implementation of the data augmentation (DA) algorithm.[20] We report multivariable risk ratios with 95% confidence intervals. Statistical analysis was performed using SAS Enterprise Guide Version 8.3 and Stata/SE 16.1.[21]

We replicated the 4C Mortality Score described by Knight et al. with one modification.[17] Since we did not have a variable for Glasgow Coma Score or confusion on exam, we used the symptom "altered mental status or confusion" instead. In addition to the full score, we tested a modified score dropping CRP, which was missing in 39% of the records. We evaluated discrimination and calibration using 9 score categories available from Knight et al.: (0-2, 3-4, 5-6, 7-8, 9-10, 11-12, 13-14, 15-16, ≥17). We used the mortality reported in the 4C validation dataset as our predicted risks for comparison with observed mortality. For reporting, we pooled the results into the 4 risk groups defined by Knight et al.: Low 0-3, Intermediate 4-8, High 9-14, and Very High ≥15. The AUROC was calculated with 95% confidence interval using the DeLong method.[22] Calibration was assessed using a standard calibration table, mean prediction error, and the square root of both the calibration error and the Brier score. We also used a modified Bland-Altman-style calibration plot.[23]

RESULTS

Of 26,914 patients in the first version of the registry, 6,822 met inclusion criteria for this analysis (≥ 18 years old, SARS-CoV-2+, hospitalized from the ED, full code status). We excluded 11 who were missing vital status at 30 days and 9 who died in the ED prior to vital signs, leaving 6,802 in the analysis cohort. Of the 6,802 , 1,149 (16.9%) died within 30 days. The median age of patients in the cohort was 64 years (interquartile range 52-75); 56.2% were male; and 61.4% had at least one comorbid condition. Of note,

the median oxygen saturation was 92% (IQR 87% to 95%) overall and 86% (IQR 76% to 93%) in those who died ($p < 0.0001$).

Table 2. Patient characteristics, clinical characteristics, and 30-day mortality of SARS-CoV-2+ patients hospitalized from the Emergency Department.

	Total	Median (IQR)	Deaths	Median (IQR)
Overall	6,802		1,149	
Key clinical measures				
Age, in years	6,802	64 (52 - 75)	1,149	74 (64 - 84)
Oxygen saturation, room air (%)	6,802	92 (87 - 95)	1,149	86 (76 - 93)
C-reactive protein (mg/dL)	4,163	10.5 (4.8 - 18.9)	643	17.3 (9.4 - 25.7)
Body Mass Index (BMI)*	6,058	28.7 (24.7 - 34)	970	28 (24.1 - 33.2)
Other vital signs				
Temperature (°C)	6,801	37.2 (36.7 - 37.9)	1,149	37.1 (36.6 - 37.8)
Heart rate (beats/minute)†	6,800	97 (84 - 110)	1,148	98 (83 - 112)
Systolic blood pressure (mm Hg)	6,802	130 (116 - 146)	1,149	127 (112 - 146)
Respiratory rate (breaths/minute)	6,797	20 (18 - 23)	1,148	20 (18 - 26)
Other blood tests				
White blood cell count (10^3/uL)	6,767	7.1 (5.3 - 9.9)	1,140	8.9 (6.2 - 12)
Hemoglobin (g/dL)	6,769	13 (12 - 14.6)	1,142	13 (11 - 14.5)
Platelets (10^3 cells/uL)	6,760	209 (160 - 275)	1,140	201.5 (149 - 270)
Sodium (mEq/L)	6,338	136 (133 - 139)	1,039	137 (133 - 142)
Potassium (mEq/L)	6,743	4.1 (3.8 - 4.6)	1,141	4.4 (4 - 5.1)
BUN (mg/dL)	6,706	18 (12 - 31)	1,131	32 (19 - 54)
Creatinine (mg/dL)	2,832	1 (0.8 - 1.4)	214	1.3 (1 - 2.1)
Total bilirubin (mg/dL)	6,124	0.6 (0.4 - 0.8)	1,051	0.6 (0.4 - 0.9)

IQR=Interquartile range

P-value for rank sum test comparison of died versus survived are p <= 0.0003 except:

† Heart rate, p=0.62

Of the demographic risk factors, age group, male sex, and residence in a nursing home/assisted living facility were the principal mortality predictors (Table 3). In the multivariable model, age 80+ years increased 30-day mortality risk by a factor of 5.79 (95% CI 4.23 - 7.34); male sex increased it by 1.14 (95% CI 1.05 - 1.28); and nursing home/assisted living facility residence increased it by 1.29 (95% CI 1.1 - 1.48). On a univariable basis, Hispanic ethnicity and smoking status were associated with lower mortality risk, but after adjusting for other variables, including the younger age of Hispanics and smokers, the risk ratio for mortality for Hispanic ethnicity was 0.96 (95% CI 0.82 - 1.1) and for smoking was 1.02 (95% CI 0.82 - 1.23).

In the univariable analysis, extreme obesity (BMI ≥ 40) did not increase risk, but after adjustment for age, sex, and other comorbidities, the risk ratio for BMI ≥ 40 was 1.44 (95% CI 1.23 - 1.64). In addition to obesity, the multivariable analysis (Table 3) showed that other comorbidities associated with increased risk of death were chronic cardiac disease, chronic pulmonary disease, liver disease as indicated by a total bilirubin ≥ 2.0 mg/dL, and kidney disease as indicated by a creatinine ≥ 1.2 mg/dL or BUN ≥ 40. Asthma and diabetes were not significant risk factors. Patients who arrived from a nursing home had an increased risk of mortality (Risk Ratio 1.29, 95% CI 1.1 - 1.48).

Table 3. Effect of patient and clinical characteristics on 30-day mortality of SARS-CoV-2+ patients hospitalized from the Emergency Department.

		Total	% of Sample	Deaths	30-day Mortality (% died)	Relative Risk	Multivariable Relative Risk (95% CI)
Overall		6,802		1,149	16.9		
In both 4C Score and Multivariable Model							
Age Group (years)							
	18 - 49	1,413	20.8	63	4.5	reference	reference
	50 - 59	1,272	18.7	108	8.5	1.90	1.66 (1.18-2.14)
	60 - 69	1,690	24.9	263	15.6	3.49	2.84 (2.09-3.58)
	70 - 79	1,302	19.1	293	22.5	5.05	4.03 (2.98-5.08)
	80+	1,125	16.5	422	37.5	8.41	5.79 (4.23-7.34)

Sex at birth	Female	2,980	43.8	466	15.6	reference	reference
	Male	3,822	56.2	683	17.9	1.14	1.17 (1.05-1.28)
Respiratory rate (breaths/min)	< 20	2,896	42.6	384	13.3	reference	reference
	20 - 29	3,282	48.3	567	17.3	1.30	1.12 (1-1.24)
	≥ 30	619	9.1	197	31.8	2.40	1.66 (1.42-1.9)
Oxygen saturation, room air (%)	≥ 92	4,017	59.1	364	9.1	reference	reference
	< 92	2,785	40.9	785	28.2	3.11	2.32 (2.06-2.58)
C-reactive protein (mg/dL)	< 5.0	1,064	25.6	64	6.0	reference	reference
	5.0 - 9.9	947	22.8	108	11.4	2.23	1.23 (1.01-1.45)
	≥ 10.0	2,152	51.7	471	21.9	4.52	1.7 (1.44-1.95)
In 4C Score (only toward comorbidity count) and Multivariable Model							
*Body Mass Index (BMI)							
	18.5 - < 40	5,227	86.3	823	15.7	reference	reference
	< 18.5	175	2.9	41	23.4	1.49	0.96 (0.72-1.2)
	≥ 40	656	10.8	106	16.2	1.03	1.44 (1.23-1.64)
*Cancer		547	8.0	88	16.1	0.95	0.81 (0.67-0.96)
*Chronic Cardiac Disease (any of below)		1,170	17.2	277	23.7	1.53	1.06 (0.93-1.19)
	Atrial Fibrillation	542	8.0	149	27.5	1.72	
	Heart Disease	382	5.6	69	18.1	1.07	
	Heart Failure	608	8.9	142	23.4	1.44	
*Chronic Pulmonary Disease (any of below)		529	7.8	112	21.2	1.28	1.06 (0.89-1.24)
	COPD	433	6.4	91	21.0	1.27	
	Bronchiectasis	17	0.3	3	17.6	1.04	
	Other Lung Disease	128	1.9	27	21.1	1.25	
	Pulmonary Fibrosis	18	0.3	5	27.8	1.65	
*Diabetes		2,079	30.6	357	17.2	1.02	0.97 (0.87-1.07)
*Liver Disease (Total Bilirubin ≥2.0)		175	2.9	52	29.7	1.77	1.56 (1.24-1.88)
*Kidney Disease (Creatinine ≥1.2 or BUN ≥40)		1,877	28.1	515	27.4	2.14	1.58 (1.42-1.74)
In 4C Score only							
Comorbidities, among 7 with *	0	2,632	38.7	300	11.4	reference	
	1	2,207	32.5	418	18.9	1.08	

		2 +	1,963	28.9	431	22.0	1.12	
	Altered mental status		957	14.1	162	16.9	1.00	
	BUN (mg/dL)	< 20	3,715	55.4	297	8.0	reference	
		20 - 39	1,771	26.4	390	22.0	2.75	
		≥ 40	1,220	18.2	444	36.4	4.56	
In Multivariable Model only								
	Race/ethnicity	White, non H/L	1,652	24.3	323	19.6	reference	reference
		Asian, non H/L	234	3.4	41	17.5	0.90	1.05 (0.77-1.33)
		Black, non H/L	2,286	33.6	362	15.8	0.81	1 (0.87-1.13)
		Hispanic/Latinx (H/L)	1,759	25.9	228	13.0	0.66	0.96 (0.82-1.1)
		Other or Unknown	871	12.8	195	22.4	1.14	1.3 (1.11-1.49)
	Resides in Nursing Home or Assisted Living		703	10.3	256	36.4	2.49	1.29 (1.1-1.48)
	Smoker		447	7.5	48	10.7	0.70	1.02 (0.82-1.23)
	Asthma		581	8.5	68	11.7	0.67	0.89 (0.71-1.06)

CI=Confidence Interval; COPD=Chronic Obstructive Pulmonary Disease
Where missing is over 1.5%: C-reactive protein=38.8%; BMI=10.9%; Total Bilirubin=10.0%; Smoker=12.7%.
"Relative Risk" is the risk of death relative to the reference if indicated, otherwise to not having the risk factor.

Table 3 also shows that increase in respiratory rate, decrease in oxygen saturation, and increase in CRP each corresponded with increase in mortality,.

Compared with the 4C validation dataset from Knight et al., the mean 4C Mortality Scores were lower in our dataset (mean score 9.0 vs. 10.6). (Figure 1A). The AUROC from the RECOVER dataset was comparable to that of the original 4C validation dataset. Using nine 4C score categories, the AUROC from the RECOVER dataset was not substantially different than the AUROC from the original 4C validation dataset (AUROC: RECOVER 0.786 (95% CI 0.773 - 0.799) vs 4C validation 0.763 [95% CI 0.757 - 0.769). (Figure 1B). Our observed category-specific mortalities were lower than those in the 4C validation dataset. Using the mortalities from the 4C validation dataset would have over-estimated risk by 6.0% on average. (Mean prediction error 6.0%. √Calibration Error 0.066 and √Brier Score 0.350). (Table 4). (Figure 1C).

Figure 1. Comparison of 4C validation and RECOVER datasets

Table 4. Comparison of observed mortality by 4C Mortality Risk Group for RECOVER dataset of SARS-CoV-2+ patients hospitalized from the Emergency Department.

4C Mortality Risk Group (Score Range)	RECOVER dataset with CRP*			RECOVER dataset without CRP		
	Mortality Predicted by 4C †	Observed Mortality	Prediction Error	Mortality Predicted by 4C †	Observed Mortality	Prediction Error
Overall	22.9%	16.9%	6.0%	17.6%	16.9%	0.7%
Low (0-3 points)	1.4%	0.8%	0.6%	1.2%	1.6%	-0.5%
Intermediate (4-8 points)	9.7%	5.3%	4.4%	9.5%	8.7%	0.8%
High (9-14 points)	29.9%	22.3%	7.6%	28.6%	27.4%	1.2%
Very High (≥15 points)	60.2%	50.6%	9.6%	56.8%	58.2%	-1.4%

	AUROC	√Calibration Error	√Brier Score
4C Validation	0.763 (0.757 - 0.769)		
RECOVER dataset with CRP	0.786 (0.773 - 0.799)	0.066	0.350
RECOVER dataset without CRP	0.776 (0.762 - 0.79)	0.017	0.346

* CRP=C-reactive protein.

† The mortality predicted by 4C is constant for each individual score, but when the scores are grouped into ranges (as they are here), the predicted mortality varies based on the proportion of patients from the test dataset with each individual score within the range.

Dropping CRP from the 4C Mortality Score reduced the scores overall (mean score 7.7) but did not substantially change discrimination (AUROC 0.776, 95% CI 0.762 - 0.790). Dropping the CRP component did affect calibration. The category-specific mortalities in our dataset were now close to those in the 4C validation dataset. Using the mortalities from the 4C validation dataset would have mis-estimated risk by 0.7% on average. (Mean prediction error 0.7%. √Calibration Error 0.017, and √Brier Score 0.346).

DISCUSSION

In this analysis of multicenter data from the RECOVER Network, our results confirmed several previous findings for risk factors for COVID-19 mortality, including older age, comorbidities, BMI ≥ 40 kg/m^2, higher respiratory rate, and lower oxygen saturation.[4–9,11–14] In addition, as reported by Graselli et al.

in critically ill patients, we observed that male sex is predictive of mortality.[7] We also observed the expected, but previously unquantified finding that arrival to the emergency care setting from a nursing home was associated with increased mortality. While this has not been specifically mentioned in other studies, Ferrando-Vivas et al. found that functional dependence was related to mortality (Hazard Ratio 1.425).[5]

In the RECOVER Network, COVID-19 related hospitalizations are higher among SARS-CoV-2+ Hispanic patients when compared to Non-Hispanics, but the adjusted mortality is similar to non-Hispanic whites.[24] Similarly, Mackey et al. reported that hospitalizations for COVID-19 among those who identify their ethnicity as Hispanic were proportionately higher than for their non-Hispanic white counterparts but the case fatality rate was similar between Hispanic and non-Hispanic patients.[25]

We also found that the comorbid conditions such as liver disease defined as elevated total bilirubin ≥2.0 and kidney disease defined as creatinine ≥1.2 mg/dl or BUN ≥40 had an independent association with 30-day mortality in hospitalized SARS-CoV-2+ patients. Surprisingly, previous studies and our results did not establish diabetes as a significant risk factor.[26-28] But our findings on the association of smoking with 30-day mortality did not concur with previous studies. Smoking as well as cumulative smoking exposure was predictive of mortality in previous studies,[26] but we did not find a statistically significant association after controlling for other variables. Finally, among the clinical variables, tachypnea (respiratory rate ≥20) and hypoxemia (oxygen saturation <92%) were significant predictors of mortality. Zhao et al reported higher odds of mortality (adjusted OR 4.8) for an oxygen saturation <92%.[13]

Given the multiplicity of variables associated with 30-day mortality, clinicians need a simple score to better predict short-term mortality. The 4C Mortality Score is one such score and it performed well in our dataset. Discrimination was excellent, and calibration was also good, although using the category-specific mortalities from the 4C validation dataset would have over-estimated risk. CRP was missing in 39% of the records in our study, so we examined the performance of the 4C Mortality Score without the CRP component. Discrimination remained good, and the category-specific risks from the 4C validation were accurate. When CRP was removed from the score, many patients with high CRP values moved into a lower risk category. Those patients who remained with high 4C Mortality Scores despite removal of CRP died at a higher rate than those whose risk score decreased, but those with high CRP values who moved to a lower risk group had higher mortality than the average for their new lower risk group. This might be referred to as stage migration effect. When the high CRP patients moved from the very-high risk group to the high-risk group, the average mortality went up in both groups. Based on our observations, we suggest

using the 4C Mortality Score without the CRP component, but recalibrating risk estimates as per our Table 4 or Supplementary Table A. Using category-specific risks as opposed to the 4 risk groups (low, intermediate, high, very high) is preferred because it doesn't assume the distribution across the risk groups is the same in different populations. This modified 4C Mortality Score could assist with triage decisions, to inform patients and their family members of prognostic information, and to help with forecasting of resource utilization in the hospital.

The nature of the COVID-19 pandemic greatly accelerated the timeline of related research and has resulted in rapid changes to practice patterns and patient presentation. At the time of this study, the 4C Mortality Score was among the most promising risk evaluation tools, and had been identified has having a low likelihood of bias. Since the inception of our study to validate this score, many others systems have been proposed. These have been developed in a variety of different patient populations using a wide range of methods.[27–35] Some models have been independently assessed and performance varies.[36] Updates to a systematic review of prediction models continue to identify the prognostic 4C Mortality Score as among the most promising [37]suggesting that attempts to validate and calibrate this and other existing risk estimation models could aid providers in the evaluation of the many available scoring systems for patients with COVID-19 disease.

Limitations

This is a national study of hospitalized SARS-CoV-2+ patients. The large sample size, the number and diversity of the participating sites, and a comprehensive list of data elements are major strengths of this study. However, some sites contributed more SARS-CoV-2+ patients than others. We did see regional differences in 30-day mortality, but these did not affect the risk ratios. As noted above, CRP was missing in almost 39% of patients.

Additional limitations are related to the nature of the COVID-19 pandemic and the changes in patient population and clinical practices that have occurred over time. The data in this study represent a time period early in the pandemic (on or before September 2020) and thus may not fully account for practice changes. However, these data align with the time period of the RECOVERY trial, which introduced the main practice change affecting mortality (use of dexamethasone) in February 2021.[38]

CONCLUSIONS

We conclude that among SARS-CoV-2+ hospitalized patients, older patients with comorbid conditions and those with hypoxemia at the time of presentation have a very high risk of dying within 30 days. We independently validate the 4C Mortality Score as predicting risk of death in hospitalized SARS-CoV-2+ patients, but we recommend dropping the CRP component of the score and using our recalibrated mortality risk estimates.

Conflict of Interests: None

Grants: Funded by Department of Emergency Medicine, Indiana University School of Medicine. No Grant Number.

Author contributions: PG, MAK, AJG, CB, CC and JK conceived and designed the study, JK obtained research funding, designed, and organized the registry. MAK, LM provided statistical advice on study design and analyzed the data; AJG, MAK, PG drafted the manuscript, and all authors contributed substantially to its revision. PG takes responsibility for the paper as a whole.

Data sharing agreement: No data is available

Ethics approval statement: The RECOVER registry protocol was reviewed by the institutional review boards (IRBs) at all sites; 3 IRBs provided approval with waiver of informed consent. All others provided an exemption from human subjects designation. All data were anonymized prior to analysis. Local IRB: 56234

References

1 Lipsitch M, Swerdlow DL, Finelli L. Defining the Epidemiology of Covid-19 — Studies Needed. New England Journal of Medicine 2020;382:1194–6.

2 Kerlin MP, Costa DK, Davis BS, et al. Actions Taken by US Hospitals to Prepare for Increased Demand for Intensive Care During the First Wave of COVID-19: A National Survey. In: LID.0012-3692 21 00478-5.

3 Yang X, Yu Y, Xu J, et al. Clinical course and outcomes of critically ill patients with SARS-CoV-2 pneumonia in Wuhan, China: a single-centered, retrospective, observational study.

4 Ferrando C, Mellado-Artigas R, Gea A, et al. Patient characteristics, clinical course and factors associated to ICU mortality in critically ill patients infected with SARS-CoV-2 in Spain: A prospective, cohort, multicentre study. Rev Esp Anestesiol Reanim 2020;67:425–37.

5 Ferrando-Vivas P, Doidge J, Thomas K, et al. Prognostic Factors for 30-Day Mortality in Critically Ill Patients With Coronavirus Disease 2019: An Observational Cohort Study. Crit Care Med 2021;49:102–11.

6 Gallo Marin B, Aghagoli G, Lavine K, et al. Predictors of COVID-19 severity: A literature review. Reviews in Medical Virology 2021;31.

7 Grasselli G, Greco M, Zanella A, et al. Risk Factors Associated With Mortality Among Patients With COVID-19 in Intensive Care Units in Lombardy, Italy. JAMA Internal Medicine 2020;180:1345–55.

8 Liang W, Liang H, Ou L, et al. Development and Validation of a Clinical Risk Score to Predict the Occurrence of Critical Illness in Hospitalized Patients With COVID-19. JAMA Internal Medicine 2020;180:1081–9.

9 Ruan Q, Yang K, Wang W, et al. Clinical predictors of mortality due to COVID-19 based on an analysis of data of 150 patients from Wuhan, China. Intensive Care Med 2020;46:846–8.

10 Wynants L, Van Calster B, Collins GS, et al. Prediction models for diagnosis and prognosis of covid-19: systematic review and critical appraisal. BMJ 2020;369:m1328.

11 Xu J, Yang X, Yang L, et al. Clinical course and predictors of 60-day mortality in 239 critically ill patients with COVID-19: a multicenter retrospective study from Wuhan, China. Crit Care 2020;24.

12 Zhang S, Guo M, Duan L, et al. Development and validation of a risk factor-based system to predict short-term survival in adult hospitalized patients with COVID-19: a multicenter, retrospective, cohort study. Crit Care 2020;24.

13 Zhao Z, Chen A, Hou W, et al. Prediction model and risk scores of ICU admission and mortality in COVID-19. PLoS One 2020;15.

14 Zhou F, Yu T, Du R, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. The Lancet 2020;395:1054–62.

15 Levy TJ, Richardson S, Coppa K, et al. A predictive model to estimate survival of hospitalized COVID-19 patients from admission data. medRxiv;2020.

16 Brajer N, Cozzi B, Gao M, et al. Prospective and External Evaluation of a Machine Learning Model to Predict In-Hospital Mortality of Adults at Time of Admission. JAMA Network Open 2020;3.

17 Knight H SR, A P, R B, et al. Risk stratification of patients admitted to hospital with covid-19 using the ISARIC WHO Clinical Characterisation Protocol: development and validation of the 4C Mortality Score. BMJ 2020;370:m3339.

18 Kline JA, Pettit KL, Kabrhel C, et al. Multicenter registry of United States emergency department patients tested for SARS-CoV-2. J Am Coll Emerg Physicians Open 2020;1:1341–8.

19 Niknam KK, Mitarai T, Nesbitt J, et al. Emergency Critical Care SOFA Score Predicts In-Hospital Mortality. Society of Academic Emergency Medicine 2019.

20 StataCorp. Stata Multiple-Imputation Reference Manual College Station. 2021.https://www.stata.com/manuals/mi.pdf.

21 Cummings P. Estimating Adjusted Risk Ratios for Matched and Unmatched Data: An Update. The Stata Journal 2011;11:290–8.

22 DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 1988;44:837–45.

23 Newman TB, Kohn MA. Evidence-based diagnosis : an introduction to clinical epidemiology. Cambridge ; New York: : Cambridge University Press 2019.

24 Kline JA, Camargo CA Jr, Courtney DM, et al. Clinical prediction rule for SARS-CoV-2 infection from 116 U.S. emergency departments 2-22-2021. PLOS ONE 2021;16.

25 Mackey K, Ayers CK, Kondo KK, et al. Racial and Ethnic Disparities in COVID-19-Related Infections, Hospitalizations, and Deaths : A Systematic Review. Ann Intern Med 2021;174:362–73.

26 Lowe KE, Zein J, Hatipoğlu U, et al. Association of Smoking and Cumulative Pack-Year Exposure With COVID-19 Outcomes in the Cleveland Clinic COVID-19 Registry. JAMA Internal Medicine 2021.

27 Cho S-Y, Park S-S, Song M-K, et al. Prognosis Score System to Predict Survival for COVID-19 Cases: a Korean Nationwide Cohort Study. J Med Internet Res 2021;23:e26257–e26257. doi:10.2196/26257

28 Berenguer J, Borobia AM, Ryan P, et al. Development and validation of a prediction model for 30-day mortality in hospitalised patients with COVID-19: the COVID-19 SEIMC score. Thorax 2021;76:920–9. doi:10.1136/thoraxjnl-2020-216001

29 Nicholson CJ, Wooster L, Sigurslid HH, et al. Estimating Risk of Mechanical Ventilation and Mortality Among Adult COVID-19 patients Admitted to Mass General Brigham: The VICE and DICE Scores. medRxiv 2020;:2020.09.14.20194670. doi:10.1101/2020.09.14.20194670

30 Magro B, Zuccaro V, Novelli L, et al. Predicting in-hospital mortality from Coronavirus Disease 2019: A simple validated app for clinical use. PLoS One 2021;16:e0245281–e0245281. doi:10.1371/journal.pone.0245281

31 King JT Jr, Yoon JS, Rentsch CT, et al. Development and validation of a 30-day mortality index based on pre-existing medical administrative data from 13,323 COVID-19 patients: The Veterans Health Administration COVID-19 (VACO) Index. PLoS One 2020;15:e0241825–e0241825. doi:10.1371/journal.pone.0241825

32 Goodacre S, Thomas B, Sutton L, et al. Derivation and validation of a clinical severity score for acutely ill adults with suspected COVID-19: The PRIEST observational cohort study. PLoS One 2021;16:e0245840–e0245840. doi:10.1371/journal.pone.0245840

33 Gupta RK, Harrison EM, Ho A, et al. Development and validation of the ISARIC 4C Deterioration model for adults hospitalised with COVID-19: a prospective cohort study. Lancet Respir Med 2021;9:349–59. doi:10.1016/S2213-2600(20)30559-2

34 Hohl CM, Rosychuk RJ, Archambault PM, et al. Derivation and validation of a clinical score to predict death among non-palliative covid-19 patients presenting to emergency departments: the CCEDRRN COVIDmortality score. medRxiv 2021;:2021.07.28.21261283. doi:10.1101/2021.07.28.21261283

35 Halasz G, Sperti M, Villani M, et al. A Machine Learning Approach for Mortality Prediction in COVID-19 Pneumonia: Development and Evaluation of the Piacenza Score. J Med Internet Res 2021;23:e29058–e29058. doi:10.2196/29058

36 Gupta RK, Marks M, Samuels THA, et al. Systematic evaluation and external validation of 22 prognostic models among hospitalised adults with COVID-19: an observational cohort study. Eur Respir J 2020;56:2003498. doi:10.1183/13993003.03498-2020

37 Update to living systematic review on prediction models for diagnosis and prognosis of covid-19. BMJ 2021;372. doi:10.1136/bmj.n236

38 The RECOVERY Collaborative Group. Dexamethasone in Hospitalized Patients with Covid-19. N Engl J Med 2021;384:693–704. doi:10.1056/NEJMoa2021436

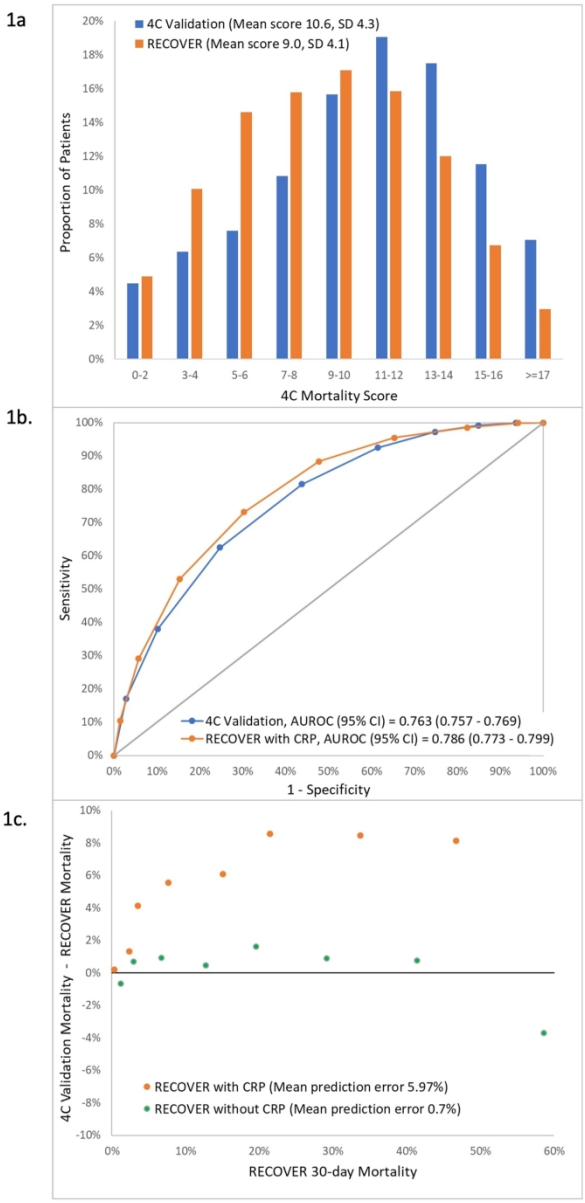

Figure 1: Comparison of 4C validation and RECOVER datasets 1a. 4C Mortality Scores were lower in the RECOVER dataset than in the original 4C validation dataset. 1b. ROC curves for the 4C Mortality Score (categorized into the 9 ranges from Figure 1a) in the 4C validation dataset and the RECOVER dataset. 1c. Calibration plot (modified Bland-Altman) showing prediction error versus observed mortality for the 4C Mortality Score with and without the C-reactive protein (CRP) component. Points from left to right are in the 4C Mortality Score ranges shown in Figure 1A from left to right.

118x238mm (300 x 300 DPI)

Table A. Observed mortality by 4C Risk Group calculated without C-reactive protein for RECOVER dataset of SARS-CoV-2+ patients hospitalized from the Emergency Department.

4C Mortality Risk Group	Patient Distribution	Predicted Risk	Observed Mortality	Prediction Error
0-2	8.1%	0.5%	1.6%	-1.1%
3-4	10.8%	3.7%	3.7%	-0.1%
5-6	14.4%	7.6%	8.4%	-0.7%
7-8	16.9%	13.2%	16.0%	-2.8%
9-10	17.9%	21.2%	25.9%	-4.7%
11-12	16.3%	30.0%	34.4%	-4.4%
13-14	10.3%	42.2%	49.9%	-7.7%
15-16	4.7%	54.9%	62.4%	-7.5%
≥17	0.7%	72.3%	66.1%	6.3%

Mean Prediction Error	-3.3%
√Calibration Error	0.0421
√Brier Score	0.3875

STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No	Recommendation
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract "retrospective cohort study" √ **(page 3)**
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found√ **(page 3)**
Introduction		
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported√**(page 6)**
Objectives	3	State specific objectives, including any prespecified hypotheses√**(page6,7)**
Methods		
Study design	4	Present key elements of study design early in the paper√ **(page6)**
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection **(page 3,7)**
Participants	6	**(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up √(page 6)** *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed *Case-control study*—For matched studies, give matching criteria and the number of controls per case
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable √ **(page 7)**
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group √ **(page 7)**
Bias	9	Describe any efforts to address potential sources of bias √ **(page 3, 12)**
Study size	10	Explain how the study size was arrived at √ **(page 7)**
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why √ **(page 7-9)**
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding √ **(page 8)**
		(*b*) Describe any methods used to examine subgroups and interactions √ **(page 8)**
		(*c*) Explain how missing data were addressed √ **(page 8)**
		(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed *Case-control study*—If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy
		(*e*) Describe any sensitivity analyses

Continued on next page

Results		
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed √
		(b) Give reasons for non-participation at each stage
		(c) Consider use of a flow diagram
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders √
		(b) Indicate number of participants with missing data for each variable of interest
		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount) 30 day mortality √
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time **√ (page 7)**
		Case-control study—Report numbers in each exposure category, or summary measures of exposure
		Cross-sectional study—Report numbers of outcome events or summary measures
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included **√ (page 9,11)**
		(*b*) Report category boundaries when continuous variables were categorized
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses **√ (page 9,10)**
Discussion		
Key results	18	Summarise key results with reference to study objectives**√ (page 9)**
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias**√(page 12)**
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence**√ (page 12)**
Generalisability	21	Discuss the generalisability (external validity) of the study results**√ (page 12)**
Other information		
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based**√ (page 2)**

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

BMJ Open

External Validation of the 4C Mortality Score for Hospitalized COVID-19 Patients in the RECOVER network

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2021-054700.R2
Article Type:	Original research
Date Submitted by the Author:	31-Jan-2022
Complete List of Authors:	Gordon, Alexandra June; Stanford University School of Medicine, Emergency Medicine Govindarajan, Prasanthi; Stanford University School of Medicine, Department of Emergency Medicine Bennett, Christopher L.; Stanford University School of Medicine, Department of Emergency Medicine Matheson, Loretta ; Stanford University School of Medicine, Department of Emergency Medicine Kohn, Michael; Stanford University School of Medicine, Emergency Medicine; University of California San Francisco, Epidemiology and Biostatistics Camargo, Carlos; Massachusetts General Hospital, Harvard Medical School, Emergency Medicine Kline, Jeffrey ; Wayne State University School of Medicine, Emergency Medicine
Primary Subject Heading:	Emergency medicine
Secondary Subject Heading:	Infectious diseases
Keywords:	COVID-19, ACCIDENT & EMERGENCY MEDICINE, Adult intensive & critical care < ANAESTHETICS, EPIDEMIOLOGY, GENERAL MEDICINE (see Internal Medicine)

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

External Validation of the 4C Mortality Score for Hospitalized COVID-19 Patients in the RECOVER network

Alexandra June Gordon, MD (First author)

Department of Emergency Medicine, Stanford University School of Medicine,

ajgordon@stanford.edu

Prasanthi Govindarajan, MBBS, MAS (Co-first author)

Department of Emergency Medicine, Stanford University School of Medicine

prashagovindarajan@stanford.edu

Christopher L. Bennett, MD

Department of Emergency Medicine, Stanford University School of Medicine

cleebennett@stanford.edu

Loretta Matheson, MS

Department of Emergency Medicine, Stanford University School of Medicine

lorettamatheson@stanford.edu

Michael A. Kohn, MD

Department of Emergency Medicine, Stanford University School of Medicine

Michael.Kohn@ucsf.edu

Carlos A. Camargo, Jr., MD, DrPH

Department of Emergency Medicine, Massachusetts General Hospital, Harvard Medical School, Boston, MA

CCAMARGO@PARTNERS.ORG

Jeffrey A. Kline MD

Department of Emergency Medicine, Wayne State University School of Medicine

jkline@wayne.edu

Corresponding author: Jeffrey Kline MD, jkline@wayne.edu

Department of Emergency Medicine,

Wayne State University
Department of Emergency Medicine
4201 St. Antoine
University Health Center - 6G
Detroit, MI 48201

Manuscript word count (not including title, abstract, acknowledgment, references, tables, and figure legends): 2707

ABSTRACT:

Objectives: Estimating mortality risk in hospitalized SARS-CoV-2+ patients may help with choosing level of care and discussions with patients. The Coronavirus Clinical Characterisation Consortium Mortality Score (4C Score) is a promising COVID-19 mortality risk model. We examined the association of risk factors with 30-day mortality in hospitalized, full-code SARS-CoV-2+ patients and investigated the discrimination and calibration of the 4C Score. This was a retrospective cohort study of SARS-CoV-2+ hospitalized patients within the RECOVER (REgistry of suspected COVID-19 in EmeRgency care) Network.

Setting: 99 emergency departments (EDs) across the US.

Participants: Patients ≥ 18 years old, positive for SARS-CoV-2 in the ED, and hospitalized.

Primary outcome: Death within 30 days of the index visit. We performed logistic regression analysis, reporting multivariable risk ratios (MVRRs) and calculated the AUROC and mean prediction error for the original 4C Score and after dropping the c-reactive protein component (CRP).

Results: Of 6,802 hospitalized COVID-19 patients, 1,149 (16.9%) died within 30 days. The 30-day mortality was increased with age 80+ years (MVRR = 5.79, 95% CI 4.23 - 7.34); male sex (MVRR = 1.17,1.05 - 1.28); and nursing home/assisted living facility residence (MVRR = 1.29, 1.1 - 1.48). The 4C Score had comparable discrimination in the RECOVER dataset compared with the original 4C validation dataset (AUROC: RECOVER 0.786 (95% CI 0.773 - 0.799), 4C validation 0.763 (95% CI 0.757 - 0.769). Score-specific mortalities in our sample were lower than in the 4C validation sample (mean prediction error 5.97%). Dropping the CRP component

from the 4C Score did not substantially affect discrimination and 4C risk estimates were now close (mean prediction error 0.7%).

Conclusions: We independently validated 4C Score as predicting risk of 30-day mortality in hospitalized SARS-CoV-2+ patients. We recommend dropping the CRP component of the score and using our recalibrated mortality risk estimates.

ARTICLE SUMMARY:

Strengths and Limitations:

- In this first study using a national US sample of patients who tested positive for SARS-CoV-2 and were hospitalized through emergency departments, our results confirmed the previous findings that older age, comorbidities, BMI ≥ 40 kg/m2, higher respiratory rate, and lower oxygen saturation were associated with 30-day mortality.
-We also observed that the arrival to the emergency care setting from a nursing home was associated with increased mortality.
-We independently validated 4C Mortality Score as predicting risk of 30-day mortality in hospitalized SARS-CoV-2+ patients.
-We recommend dropping the CRP component of the score and using our recalibrated mortality risk estimates when estimating the 30-day mortality in hospitalized patients who test positive for SARS-CoV-2.

INTRODUCTION

The COVID-19 pandemic has placed tremendous strain on emergency and critical care resources in hospitals worldwide.[1–3] To prepare the healthcare systems for the surges, several studies have developed prediction models to assess mortality risk in patients hospitalized with COVID-19. These studies identified the following risk factors for mortality or critical care admission: age, sex, comorbid conditions, laboratory values, and vital signs.[4–16]

In a systematic review that evaluated many of these risk prediction studies using the prediction model risk of bias assessment tool (PROBAST), Wynants et al. concluded that many of the current risk models may be misleading.[10] However, the authors' analysis suggested that one COVID-19 mortality prediction model, the 4C Mortality Score, which was built on a large UK data set, had relatively low risk of bias in most domains by the PROBAST criteria. The 4C Mortality Score includes eight variables - age, sex, respiratory rate, oxygen saturation, number of comorbidities, level of consciousness, blood urea nitrogen, and C-reactive protein (CRP); see table 1.[17] While there has been continued interest in the development of prediction models for COVID-19, the 4C Mortality Score represented one of the first with a low risk of bias and therefore a good candidate for verification in other populations.

Table 1. Point assignment for 4C Mortality Score

Age Group (years)	18 - 49	
	50 - 59	+ 2
	60 - 69	+ 4
	70 - 79	+ 6

	80+	+ 7
Sex at birth	Female	
	Male	+ 1
Comorbidities*	0	
	1	+ 1
	2 +	+ 2
Respiratory rate (breaths/minute)	< 20	
	20 - 29	+ 1
	≥ 30	+ 2
Oxygen saturation, room air (%)	≥ 92	
	< 92	+ 2
Altered mental status †	No	
	Yes	+ 2
BUN (mg/dL)§	< 20	
	20 - 39	+ 1
	≥ 40	+ 3
C-reactive protein (mg/dL)§	< 5.0	
	5.0 - 9.9	+ 1
	≥ 10.0	+ 2

* Comorbidities: High Body Mass Index, Cancer, Chronic cardiac disease, Chronic pulmonary disease, Diabetes, Liver disease, Kidney disease.

† Altered mental status is patient-reported symptom whereas 4C used Glasgow coma scale < 15.

§ BUN and C-reactive protein units converted from 4C.

In this study, we investigated the risk of 30-day mortality in hospitalized SARS-CoV-2+ patients within the RECOVER (REgistry of suspected COVID-19 in EmeRgency care) Network.[18] In a large cohort of SARS-CoV-2+ patients hospitalized from 99 US emergency departments (EDs), we determined the relation of demographic and clinical factors with 30-day mortality and investigated the discrimination and calibration of the 4C Mortality Score with and without the CRP value.

METHODS

In this retrospective cohort study, we included patient-level data from the RECOVER Network, a national registry of patients who were tested for SARS-CoV-2 during their ED visit. We restricted the analysis to full code status patients ≥ 18 years old who tested positive for SARS-CoV-2 and were hospitalized from the ED.[18] The study was approved or deemed exempted by the Institutional Review Boards of all participating sites.

Data Source

We obtained data from 40 medical centers representing 99 EDs from 27 US states and the District of Columbia. Data were collected using a REDCap data collection form that was distributed to the ED sites during the study period (3/2020-9/2020); our data were downloaded from the registry in 12/2020. The REDCap form (Appendix A) had seven sections and 204 questions, which generated 360 data elements. Variables reflected a combination of routinely collected information (e.g. patient demographics, medical history, vital signs, and diagnostic test results), patient-reported symptoms and risk exposures, clinical outcomes (e.g. admission, therapies, death), and those deemed important by the RECOVER Network steering committee. After creation, but prior to launch, the data form was piloted at 19 sites and refined. For additional section details and the questions, please refer to the data collection form in the appendix (supplement). The data were obtained from the electronic healthcare record using a combination of electronic download for routinely collected, coded variables (e.g., age, vital signs

and laboratory values), supplemented by chart review by research personnel, using methods previously described.[18]

Patient and Public Involvement

Patients were not involved in the development of our work, setting the research question or determining the outcome measures. This applies to both the RECOVER Network and the work presented here. Given the nature and limitations of emergency care during the COVID-19 pandemic, it was not appropriate or possible to involve patients or the public in the design, or conduct, reporting, or dissemination plans of our work.

Study Variables

We analyzed patient characteristics such as demographics, vital signs, symptoms, risks for infection, comorbidities, and medications. Following the 4C Mortality Score, we categorized the patients into five age groups (18-49, 50-59, 60-69, 70-79, 80+). The US standard ethnicity (Latinx/Hispanic yes/no) and race categories were combined into 8 categories (Non-Hispanic White, Non-Hispanic African American/Black, Non-Hispanic Asian, Non-Hispanic Native American/Alaskan Native, Non-Hispanic Pacific Islander/Native Hawaiian, Non-Hispanic mixed, and Non-Hispanic Unknown). In the analysis, we combined Non-Hispanic Native American (0.2%), Non-Hispanic Pacific Islander (0.2%), Non-Hispanic mixed, Other, and Unknown into a single group (12.8%).

All included patients had a positive reverse transcriptase polymerase chain reaction test (RT-PCR) test for SARS-CoV-2. Almost all the tests were performed during the ED visit, but we also included patients who had a test in a physician's office or urgent care immediately prior to the ED visit. We excluded patients whose 30-day vital status could not be ascertained, those who died in the ED before vital signs were recorded, and those who did not have full code status.

Study Outcome

The primary outcome was death within 30 days of the index visit. The 4C mortality score's predictive accuracy was measured by the area under the ROC curve (AUROC) and mean prediction error.

Data Preparation and Statistical Analysis

For comparison with the other cohorts, we report the median and interquartile range of continuous variables – both in the entire cohort and in the subgroup who died within 30 days – and compared the median values using the rank sum test. We performed univariable analysis on 26 independent variables that were included on the data collection form using complete (non-missing) data and reporting risk ratios for 30-day mortality. For risk ratio reporting of continuous variables, we chose category boundaries based on the 4C Mortality Score (age, respiratory rate, oxygen saturation, BUN, and CRP) or other published mortality prediction models (BMI, creatinine, total bilirubin).[19]

We selected variables for our multivariable logistic regression model based on the 4C Mortality Score, other prior studies, and clinical judgement. The RECOVER dataset has complete data (<

1.5% missing) on most variables, with the exception of CRP, body mass index (BMI), bilirubin, and smoking status. For our multivariable analysis, we used imputed values for missing data using Stata's implementation of the data augmentation (DA) algorithm.[20] We report multivariable risk ratios with 95% confidence intervals. Statistical analysis was performed using SAS Enterprise Guide Version 8.3 and Stata/SE 16.1.[21]

We replicated the 4C Mortality Score described by Knight et al. with one modification.[17] Since we did not have a variable for Glasgow Coma Score or confusion on exam, we used the symptom "altered mental status or confusion" instead. In addition to the full score, we tested a modified score dropping CRP, which was missing in 39% of the records. We evaluated discrimination and calibration using 9 score categories available from Knight et al.: (0-2, 3-4, 5-6, 7-8, 9-10, 11-12, 13-14, 15-16, ≥17). We used the mortality reported in the 4C validation dataset as our predicted risks for comparison with observed mortality. For reporting, we pooled the results into the 4 risk groups defined by Knight et al.: Low 0-3, Intermediate 4-8, High 9-14, and Very High ≥15. The AUROC was calculated with 95% confidence interval using the DeLong method.[22] Calibration was assessed using a standard calibration table, mean prediction error, and the square root of both the calibration error and the Brier score. We also used a modified Bland-Altman-style calibration plot.[23]

RESULTS

Of 26,914 patients in the first version of the registry, 6,822 met inclusion criteria for this analysis (≥ 18 years old, SARS-CoV-2+, hospitalized from the ED, full code status). We excluded 11 who were missing vital status at 30 days and 9 who died in the ED prior to vital signs, leaving 6,802 in the analysis cohort. Of the 6,802 , 1,149 (16.9%) died within 30 days. The median age of patients in the cohort was 64 years (interquartile range 52-75); 56.2% were male; and 61.4% had at least one comorbid condition. Of note, the median oxygen saturation was 92% (IQR 87% to 95%) overall and 86% (IQR 76% to 93%) in those who died ($p < 0.0001$). (Table 2)

Table 2. Patient characteristics, clinical characteristics, and 30-day mortality of SARS-CoV-2+ patients hospitalized from the Emergency Department.

	Total	Median (IQR)	Deaths	Median (IQR)
Overall	6,802		1,149	
Key clinical measures				
Age, in years	6,802	64 (52 - 75)	1,149	74 (64 - 84)
Oxygen saturation, room air (%)	6,802	92 (87 - 95)	1,149	86 (76 - 93)
C-reactive protein (mg/dL)	4,163	10.5 (4.8 - 18.9)	643	17.3 (9.4 - 25.7)
Body Mass Index (BMI)*	6,058	28.7 (24.7 - 34)	970	28 (24.1 - 33.2)
Other vital signs				
Temperature (°C)	6,801	37.2 (36.7 - 37.9)	1,149	37.1 (36.6 - 37.8)
Heart rate (beats/minute)†	6,800	97 (84 - 110)	1,148	98 (83 - 112)
Systolic blood pressure (mm Hg)	6,802	130 (116 - 146)	1,149	127 (112 - 146)
Respiratory rate (breaths/minute)	6,797	20 (18 - 23)	1,148	20 (18 - 26)
Other blood tests				
White blood cell count (10^3/uL)	6,767	7.1 (5.3 - 9.9)	1,140	8.9 (6.2 - 12)
Hemoglobin (g/dL)	6,769	13 (12 - 14.6)	1,142	13 (11 - 14.5)

Platelets (10^3 cells/uL)	6,760	209 (160 - 275)	1,140	201.5 (149 - 270)
Sodium (mEq/L)	6,338	136 (133 - 139)	1,039	137 (133 - 142)
Potassium (mEq/L)	6,743	4.1 (3.8 - 4.6)	1,141	4.4 (4 - 5.1)
BUN (mg/dL)	6,706	18 (12 - 31)	1,131	32 (19 - 54)
Creatinine (mg/dL)	2,832	1 (0.8 - 1.4)	214	1.3 (1 - 2.1)
Total bilirubin (mg/dL)	6,124	0.6 (0.4 - 0.8)	1,051	0.6 (0.4 - 0.9)

IQR=Interquartile range

P-value for rank sum test comparison of died versus survived are p <= 0.0003 except:

† Heart rate, p=0.62

Of the demographic risk factors, age group, male sex, and residence in a nursing home/assisted living facility were the principal mortality predictors (Table 3). In the multivariable model, age 80+ years increased 30-day mortality risk by a factor of 5.79 (95% CI 4.23 - 7.34); male sex increased it by 1.14 (95% CI 1.05 - 1.28); and nursing home/assisted living facility residence increased it by 1.29 (95% CI 1.1 - 1.48). On a univariable basis, Hispanic ethnicity and smoking status were associated with lower mortality risk, but after adjusting for other variables, including the younger age of Hispanics and smokers, the risk ratio for mortality for Hispanic ethnicity was 0.96 (95% CI 0.82 - 1.1) and for smoking was 1.02 (95% CI 0.82 - 1.23).

In the univariable analysis, extreme obesity (BMI ≥ 40) did not increase risk, but after adjustment for age, sex, and other comorbidities, the risk ratio for BMI ≥ 40 was 1.44 (95% CI 1.23 - 1.64). In addition to obesity, the multivariable analysis (Table 3) showed that other comorbidities associated with increased risk of death were chronic cardiac disease, chronic

pulmonary disease, liver disease as indicated by a total bilirubin ≥ 2.0 mg/dL, and kidney disease as indicated by a creatinine ≥ 1.2 mg/dL or BUN ≥ 40. Asthma and diabetes were not significant risk factors. Patients who arrived from a nursing home had an increased risk of mortality (Risk Ratio 1.29, 95% CI 1.1 - 1.48).

Table 3. Effect of patient and clinical characteristics on 30-day mortality of SARS-CoV-2+ patients hospitalized from the Emergency Department.

		Total	% of Sample	Deaths	30-day Mortality (% died)	Relative Risk	Multivariable Relative Risk (95% CI)
Overall		6,802		1,149	16.9		
In both 4C Score and Multivariable Model							
Age Group (years)	18 - 49	1,413	20.8	63	4.5	reference	reference
	50 - 59	1,272	18.7	108	8.5	1.90	1.66 (1.18-2.14)
	60 - 69	1,690	24.9	263	15.6	3.49	2.84 (2.09-3.58)
	70 - 79	1,302	19.1	293	22.5	5.05	4.03 (2.98-5.08)
	80+	1,125	16.5	422	37.5	8.41	5.79 (4.23-7.34)
Sex at birth	Female	2,980	43.8	466	15.6	reference	reference
	Male	3,822	56.2	683	17.9	1.14	1.17 (1.05-1.28)
Respiratory rate (breaths/min)	< 20	2,896	42.6	384	13.3	reference	reference
	20 - 29	3,282	48.3	567	17.3	1.30	1.12 (1-1.24)
	≥ 30	619	9.1	197	31.8	2.40	1.66 (1.42-1.9)
Oxygen saturation, room air (%)	≥ 92	4,017	59.1	364	9.1	reference	reference
	< 92	2,785	40.9	785	28.2	3.11	2.32 (2.06-2.58)
C-reactive protein (mg/dL)	< 5.0	1,064	25.6	64	6.0	reference	reference
	5.0 - 9.9	947	22.8	108	11.4	2.23	1.23 (1.01-1.45)
	≥ 10.0	2,152	51.7	471	21.9	4.52	1.7 (1.44-1.95)
In 4C Score (only toward comorbidity count) and Multivariable Model							
*Body Mass Index (BMI)	18.5 - < 40	5,227	86.3	823	15.7	reference	reference

		< 18.5	175	2.9	41	23.4	1.49	0.96 (0.72-1.2)
		≥ 40	656	10.8	106	16.2	1.03	1.44 (1.23-1.64)
	*Cancer		547	8.0	88	16.1	0.95	0.81 (0.67-0.96)
	*Chronic Cardiac Disease (any of below)		1,170	17.2	277	23.7	1.53	1.06 (0.93-1.19)
		Atrial Fibrillation	542	8.0	149	27.5	1.72	
		Heart Disease	382	5.6	69	18.1	1.07	
		Heart Failure	608	8.9	142	23.4	1.44	
	*Chronic Pulmonary Disease (any of below)		529	7.8	112	21.2	1.28	1.06 (0.89-1.24)
		COPD	433	6.4	91	21.0	1.27	
		Bronchiectasis	17	0.3	3	17.6	1.04	
		Other Lung Disease	128	1.9	27	21.1	1.25	
		Pulmonary Fibrosis	18	0.3	5	27.8	1.65	
	*Diabetes		2,079	30.6	357	17.2	1.02	0.97 (0.87-1.07)
	*Liver Disease (Total Bilirubin ≥2.0)		175	2.9	52	29.7	1.77	1.56 (1.24-1.88)
	*Kidney Disease (Creatinine ≥1.2 or BUN ≥40)		1,877	28.1	515	27.4	2.14	1.58 (1.42-1.74)
In 4C Score only								
	Comorbidities, among 7 with *	0	2,632	38.7	300	11.4	reference	
		1	2,207	32.5	418	18.9	1.08	
		2 +	1,963	28.9	431	22.0	1.12	
	Altered mental status		957	14.1	162	16.9	1.00	
	BUN (mg/dL)	< 20	3,715	55.4	297	8.0	reference	
		20 - 39	1,771	26.4	390	22.0	2.75	
		≥ 40	1,220	18.2	444	36.4	4.56	
In Multivariable Model only								
	Race/ethnicity	White, non H/L	1,652	24.3	323	19.6	reference	reference
		Asian, non H/L	234	3.4	41	17.5	0.90	1.05 (0.77-1.33)
		Black, non H/L	2,286	33.6	362	15.8	0.81	1 (0.87-1.13)
		Hispanic/Latinx (H/L)	1,759	25.9	228	13.0	0.66	0.96 (0.82-1.1)
		Other or Unknown	871	12.8	195	22.4	1.14	1.3 (1.11-1.49)
	Resides in Nursing Home or Assisted Living		703	10.3	256	36.4	2.49	1.29 (1.1-1.48)
	Smoker		447	7.5	48	10.7	0.70	1.02 (0.82-1.23)

Asthma	581	8.5	68	11.7	0.67	0.89 (0.71-1.06)

CI=Confidence Interval; COPD=Chronic Obstructive Pulmonary Disease
Where missing is over 1.5%: C-reactive protein=38.8%; BMI=10.9%; Total Bilirubin=10.0%; Smoker=12.7%.
"Relative Risk" is the risk of death relative to the reference if indicated, otherwise to not having the risk factor.

Table 3 also shows that increase in respiratory rate, decrease in oxygen saturation, and increase in CRP each corresponded with increase in mortality,.

Compared with the 4C validation dataset from Knight et al., the mean 4C Mortality Scores were lower in our dataset (mean score 9.0 vs. 10.6). (Figure 1A). The AUROC from the RECOVER dataset was comparable to that of the original 4C validation dataset. Using nine 4C score categories, the AUROC from the RECOVER dataset was not substantially different than the AUROC from the original 4C validation dataset (AUROC: RECOVER 0.786 (95% CI 0.773 - 0.799) vs 4C validation 0.763 [95% CI 0.757 - 0.769). (Figure 1B). Our observed category-specific mortalities were lower than those in the 4C validation dataset. Using the mortalities from the 4C validation dataset would have over-estimated risk by 6.0% on average. (Mean prediction error 6.0%. √Calibration Error 0.066 and √Brier Score 0.350). (Table 4). (Figure 1C).

Figure 1. Comparison of 4C validation and RECOVER datasets

Table 4. Comparison of observed mortality by 4C Mortality Risk Group for RECOVER dataset of SARS-CoV-2+ patients hospitalized from the Emergency Department.

4C Mortality Risk Group	RECOVER dataset with CRP*	RECOVER dataset without CRP

(Score Range)	Mortality Predicted by 4C †	Observed Mortality	Prediction Error	Mortality Predicted by 4C †	Observed Mortality	Prediction Error
Overall	22.9%	16.9%	6.0%	17.6%	16.9%	0.7%
Low (0-3 points)	1.4%	0.8%	0.6%	1.2%	1.6%	-0.5%
Intermediate (4-8 points)	9.7%	5.3%	4.4%	9.5%	8.7%	0.8%
High (9-14 points)	29.9%	22.3%	7.6%	28.6%	27.4%	1.2%
Very High (≥15 points)	60.2%	50.6%	9.6%	56.8%	58.2%	-1.4%

	AUROC	√Calibration Error	√Brier Score
4C Validation	0.763 (0.757 - 0.769)		
RECOVER dataset with CRP	0.786 (0.773 - 0.799)	0.066	0.350
RECOVER dataset without CRP	0.776 (0.762 - 0.79)	0.017	0.346

* CRP=C-reactive protein.

† The mortality predicted by 4C is constant for each individual score, but when the scores are grouped into ranges (as they are here), the predicted mortality varies based on the proportion of patients from the test dataset with each individual score within the range.

Dropping CRP from the 4C Mortality Score reduced the scores overall (mean score 7.7) but did not substantially change discrimination (AUROC 0.776, 95% CI 0.762 - 0.790). Dropping the CRP component did affect calibration. The category-specific mortalities in our dataset were now close to those in the 4C validation dataset. Using the mortalities from the 4C validation dataset would have mis-estimated risk by 0.7% on average. (Mean prediction error 0.7%. √Calibration Error 0.017, and √Brier Score 0.346).

DISCUSSION

In this analysis of multicenter data from the RECOVER Network, our results confirmed several previous findings for risk factors for COVID-19 mortality, including older age, comorbidities, BMI ≥ 40 kg/m^2, higher respiratory rate, and lower oxygen saturation.[4–9,11–14] In addition, as reported by Graselli et al. in critically ill patients, we observed that male sex is predictive of mortality.[7] We also observed the expected, but previously unquantified finding that arrival to the emergency care setting from a nursing home was associated with increased mortality. While this has not been specifically mentioned in other studies, Ferrando-Vivas et al. found that functional dependence was related to mortality (Hazard Ratio 1.425).[5]

In the RECOVER Network, COVID-19 related hospitalizations are higher among SARS-CoV-2+ Hispanic patients when compared to Non-Hispanics, but the adjusted mortality is similar to non-Hispanic whites.[24] Similarly, Mackey et al. reported that hospitalizations for COVID-19 among those who identify their ethnicity as Hispanic were proportionately higher than for their non-Hispanic white counterparts but the case fatality rate was similar between Hispanic and non-Hispanic patients.[25]

We also found that the comorbid conditions such as liver disease defined as elevated total bilirubin ≥2.0 and kidney disease defined as creatinine ≥1.2 mg/dl or BUN ≥40 had an independent association with 30-day mortality in hospitalized SARS-CoV-2+ patients. Surprisingly, previous studies and our results did not establish diabetes as a significant risk factor.[26-28] But our findings on the association of smoking with 30-day mortality did not concur with previous studies. Smoking as well as cumulative smoking exposure was predictive of mortality in previous studies,[26] but we did not find a statistically significant association after

controlling for other variables. Finally, among the clinical variables, tachypnea (respiratory rate ≥20) and hypoxemia (oxygen saturation <92%) were significant predictors of mortality. Zhao et al reported higher odds of mortality (adjusted OR 4.8) for an oxygen saturation <92%.[13]

Given the multiplicity of variables associated with 30-day mortality, clinicians need a simple score to better predict short-term mortality. The 4C Mortality Score is one such score and it performed well in our dataset. Discrimination was excellent, and calibration was also good, although using the category-specific mortalities from the 4C validation dataset would have over-estimated risk. CRP was missing in 39% of the records in our study, so we examined the performance of the 4C Mortality Score without the CRP component. Discrimination remained good, and the category-specific risks from the 4C validation were accurate. When CRP was removed from the score, many patients with high CRP values moved into a lower risk category. Those patients who remained with high 4C Mortality Scores despite removal of CRP died at a higher rate than those whose risk score decreased, but those with high CRP values who moved to a lower risk group had higher mortality than the average for their new lower risk group. This might be referred to as stage migration effect. When the high CRP patients moved from the very-high risk group to the high-risk group, the average mortality went up in both groups. Based on our observations, we suggest using the 4C Mortality Score without the CRP component, but recalibrating risk estimates as per our Table 4 or Supplementary Table A. Using category-specific risks as opposed to the 4 risk groups (low, intermediate, high, very high) is preferred because it doesn't assume the distribution across the risk groups is the same in different populations. This modified 4C Mortality Score could assist with triage decisions, to inform

patients and their family members of prognostic information, and to help with forecasting of resource utilization in the hospital.

The nature of the COVID-19 pandemic greatly accelerated the timeline of related research and has resulted in rapid changes to practice patterns and patient presentation. At the time of this study, the 4C Mortality Score was among the most promising risk evaluation tools and had been identified as having a low likelihood of bias. Since the inception of our study to validate this score, many other systems have been proposed. These have been developed in a variety of different patient populations using a wide range of methods.[27–35] Some models have been independently assessed and performance varies.[36] Updates to a systematic review of prediction models continue to identify the prognostic 4C Mortality Score as among the most promising [37] suggesting that attempts to validate and calibrate this and other existing risk estimation models could aid providers in the evaluation of the many available scoring systems for patients with COVID-19 disease.

Limitations

This is a national study of hospitalized SARS-CoV-2+ patients. The large sample size, the number and diversity of the participating sites, and a comprehensive list of data elements are major strengths of this study. However, some sites contributed more SARS-CoV-2+ patients than others. We did see regional differences in 30-day mortality, but these did not affect the risk ratios. As noted above, CRP was missing in almost 39% of patients.

Additional limitations are related to the nature of the COVID-19 pandemic and the changes in patient population and clinical practices that have occurred over time. The data in this study represent a time period early in the pandemic (on or before September 2020) and thus may not fully account for practice changes. However, these data align with the time period of the RECOVERY trial, which introduced the main practice change affecting mortality (use of dexamethasone) in February 2021.[38]

CONCLUSIONS

We conclude that among SARS-CoV-2+ hospitalized patients, older patients with comorbid conditions and those with hypoxemia at the time of presentation have a very high risk of dying within 30 days. We independently validate the 4C Mortality Score as predicting risk of death in hospitalized SARS-CoV-2+ patients, but we recommend dropping the CRP component of the score and using our recalibrated mortality risk estimates.

Conflict of Interests: None

Grants: Funded by Department of Emergency Medicine, Indiana University School of Medicine. No Grant Number.

Author contributions: PG, MAK, AJG, CB, CC and JK conceived and designed the study, JK obtained research funding, designed, and organized the registry. MAK, LM provided statistical advice on study design and analyzed the data; AJG, MAK, PG drafted the manuscript, and all authors contributed substantially to its revision. PG takes responsibility for the paper as a whole.

Data sharing agreement: No data is available

Ethics approval statement: The RECOVER registry protocol was reviewed by the institutional review boards (IRBs) at all sites; 3 IRBs provided approval with waiver of informed consent. All others provided an exemption from human subjects designation. All data were anonymized prior to analysis. Local IRB: 56234

Acknowledgements:

The RECOVER Investigators Group:

Thomas Aufderheide, MD - Medical College of Wisconsin, Milwaukee, WI

Joshua Baugh, MD - Massachusetts General Hospital, Boston, MA

David Beiser, MD - University of Chicago School of Medicine, Chicago, IL

Christopher Bennett, MD - Stanford University School of Medicine, Palo Alto, CA

Joseph Bledsoe, MD - Healthcare Delivery Institute, Intermountain Healthcare, Salt Lake City, UT

Carlos A. Camargo, MD, DrPH - Massachusetts General Hospital, Harvard Medical School, Boston, MA

Edward Castillo, PhD - University of California at San Diego, San Diego, CA

Makini Chisholm-Staker, MD - Mt. Sinai School of Medicine, New York, NY

D. Mark, Courtney, MD - University of Texas Southwestern, Dallas, TX

Elizabeth Goldberg, MD - Warren Alpert Medical School of Brown University, Providence, RI

Hans House, MD - University of Iowa School of Medicine, Iowa City, IA

Stacey House, MD - Washington University School of Medicine, St. Louise, MO

Timothy Jang, MD - David Geffen School of Medicine at UCLA, Los Angeles, CA

Christopher Kabrhel, MD - Massachusetts General Hospital, Boston, MA

Jeffrey A. Kline, MD- Indiana University School of Medicine, Indianapolis, IN

Stephen Lim, MD - Louisiana State University School of Medicine, New Orleans, New Orleans, LA

Troy Madsen, MD - University of Utah School of Medicine, Salt Lake City, UT

Danielle McCarthy, MD - Northwestern University, Feinberg School of Medicine, Chicago, IL
Andrew Meltzer, MD - George Washington University School of Medicine, Washington, DC, DC
Stephen Moore, MD - Penn State Milton S. Hershey Medical Center, Hershey, PA
Mark B. Mycyk, MD- Cook County Hospitals, Chicago, IL
Craig Newgard, MD - Oregon Health and Science University, Portland, OR
Kristen E. Nordenholz, MD - University of Colorado School of Medicine, Aurora, CO
Justine Pagenhardt, MD - West Virginia University School of Medicine, Morgantown, WV
Ithan Peltan, MD - Intermountain Healthcare, Salt Lake City, UT
Katherine L. Pettit, MD - Indiana University School of Medicine, Indianapolis, IN
Michael Pulia, MD - University of Wisconsin School of Medicine and Public Health, Madison, WI
Michael Puskarich, MD - Hennepin County Medical Center and the University of Minnesota, Minneapolis, MN
Lauren Southerland, MD - Ohio State University Medical Center, Columbus, OH
Scott Sparks, MD - Riverside Regional Medical Center, Newport News, VA
Danielle Turner-Lawrence, MD - Beaumont Health, Royal Oak, MI
Marie Vrablik, MD - University of Washington School of Medicine, Seattle, WA
Alfred Wang, MD - Indiana University School of Medicine, Indianapolis, IN
Anthony Weekes, MD - Atrium Health, Charlotte, NC
Lauren Westafer, MD - Baystate Health, Springfield, MA
John Wilburn, MD - Wayne State University School of Medicine, Detroit, MI

References

1 Lipsitch M, Swerdlow DL, Finelli L. Defining the Epidemiology of Covid-19 — Studies Needed. New England Journal of Medicine 2020;382:1194–6.

2 Kerlin MP, Costa DK, Davis BS, et al. Actions Taken by US Hospitals to Prepare for Increased Demand for Intensive Care During the First Wave of COVID-19: A National Survey. In: LID.0012-3692 21 00478-5.

3 Yang X, Yu Y, Xu J, et al. Clinical course and outcomes of critically ill patients with SARS-CoV-2 pneumonia in Wuhan, China: a single-centered, retrospective, observational study.

4 Ferrando C, Mellado-Artigas R, Gea A, et al. Patient characteristics, clinical course and factors associated to ICU mortality in critically ill patients infected with SARS-CoV-2 in Spain: A prospective, cohort, multicentre study. Rev Esp Anestesiol Reanim 2020;67:425–37.

5 Ferrando-Vivas P, Doidge J, Thomas K, et al. Prognostic Factors for 30-Day Mortality in Critically Ill Patients With Coronavirus Disease 2019: An Observational Cohort Study. Crit Care Med 2021;49:102–11.

6 Gallo Marin B, Aghagoli G, Lavine K, et al. Predictors of COVID-19 severity: A literature review. Reviews in Medical Virology 2021;31.

7 Grasselli G, Greco M, Zanella A, et al. Risk Factors Associated With Mortality Among Patients With COVID-19 in Intensive Care Units in Lombardy, Italy. JAMA Internal Medicine 2020;180:1345–55.

8 Liang W, Liang H, Ou L, et al. Development and Validation of a Clinical Risk Score to Predict the Occurrence of Critical Illness in Hospitalized Patients With COVID-19. JAMA Internal Medicine 2020;180:1081–9.

9 Ruan Q, Yang K, Wang W, et al. Clinical predictors of mortality due to COVID-19 based on an analysis of data of 150 patients from Wuhan, China. Intensive Care Med 2020;46:846–8.

10 Wynants L, Van Calster B, Collins GS, et al. Prediction models for diagnosis and prognosis of covid-19: systematic review and critical appraisal. BMJ 2020;369:m1328.

11 Xu J, Yang X, Yang L, et al. Clinical course and predictors of 60-day mortality in 239 critically ill patients with COVID-19: a multicenter retrospective study from Wuhan, China. Crit Care 2020;24.

12 Zhang S, Guo M, Duan L, et al. Development and validation of a risk factor-based system to predict short-term survival in adult hospitalized patients with COVID-19: a multicenter, retrospective, cohort study. Crit Care 2020;24.

13 Zhao Z, Chen A, Hou W, et al. Prediction model and risk scores of ICU admission and mortality in COVID-19. PLoS One 2020;15.

14 Zhou F, Yu T, Du R, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. The Lancet 2020;395:1054–62.

15 Levy TJ, Richardson S, Coppa K, et al. A predictive model to estimate survival of hospitalized COVID-19 patients from admission data. medRxiv;2020.

16 Brajer N, Cozzi B, Gao M, et al. Prospective and External Evaluation of a Machine Learning Model to Predict In-Hospital Mortality of Adults at Time of Admission. JAMA Network Open 2020;3.

17 Knight H SR, A P, R B, et al. Risk stratification of patients admitted to hospital with covid-19 using the ISARIC WHO Clinical Characterisation Protocol: development and validation of the 4C Mortality Score. BMJ 2020;370:m3339.

18 Kline JA, Pettit KL, Kabrhel C, et al. Multicenter registry of United States emergency department patients tested for SARS-CoV-2. J Am Coll Emerg Physicians Open 2020;1:1341–8.

19 Niknam KK, Mitarai T, Nesbitt J, et al. Emergency Critical Care SOFA Score Predicts In-Hospital Mortality. Society of Academic Emergency Medicine 2019.

20 StataCorp. Stata Multiple-Imputation Reference Manual College Station. 2021.https://www.stata.com/manuals/mi.pdf.

21 Cummings P. Estimating Adjusted Risk Ratios for Matched and Unmatched Data: An Update. The Stata Journal 2011;11:290–8.

22 DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 1988;44:837–45.

23 Newman TB, Kohn MA. Evidence-based diagnosis : an introduction to clinical epidemiology. Cambridge ; New York: : Cambridge University Press 2019.

24 Kline JA, Camargo CA Jr, Courtney DM, et al. Clinical prediction rule for SARS-CoV-2 infection from 116 U.S. emergency departments 2-22-2021. PLOS ONE 2021;16.

25 Mackey K, Ayers CK, Kondo KK, et al. Racial and Ethnic Disparities in COVID-19-Related Infections, Hospitalizations, and Deaths : A Systematic Review. Ann Intern Med 2021;174:362–73.

26 Lowe KE, Zein J, Hatipoğlu U, et al. Association of Smoking and Cumulative Pack-Year Exposure With COVID-19 Outcomes in the Cleveland Clinic COVID-19 Registry. JAMA Internal Medicine 2021.

27 Cho S-Y, Park S-S, Song M-K, et al. Prognosis Score System to Predict Survival for COVID-19 Cases: a Korean Nationwide Cohort Study. J Med Internet Res 2021;23:e26257–e26257. doi:10.2196/26257

28 Berenguer J, Borobia AM, Ryan P, et al. Development and validation of a prediction model for 30-day mortality in hospitalised patients with COVID-19: the COVID-19 SEIMC score. Thorax 2021;76:920–9. doi:10.1136/thoraxjnl-2020-216001

29 Nicholson CJ, Wooster L, Sigurslid HH, et al. Estimating Risk of Mechanical Ventilation and Mortality Among Adult COVID-19 patients Admitted to Mass General Brigham: The VICE and DICE Scores. medRxiv 2020;:2020.09.14.20194670. doi:10.1101/2020.09.14.20194670

30 Magro B, Zuccaro V, Novelli L, et al. Predicting in-hospital mortality from Coronavirus Disease 2019: A simple validated app for clinical use. PLoS One 2021;16:e0245281–e0245281. doi:10.1371/journal.pone.0245281

31 King JT Jr, Yoon JS, Rentsch CT, et al. Development and validation of a 30-day mortality index based on pre-existing medical administrative data from 13,323 COVID-19 patients: The Veterans Health Administration COVID-19 (VACO) Index. PLoS One 2020;15:e0241825–e0241825. doi:10.1371/journal.pone.0241825

32 Goodacre S, Thomas B, Sutton L, et al. Derivation and validation of a clinical severity score for acutely ill adults with suspected COVID-19: The PRIEST observational cohort study. PLoS One 2021;16:e0245840–e0245840. doi:10.1371/journal.pone.0245840

33 Gupta RK, Harrison EM, Ho A, et al. Development and validation of the ISARIC 4C Deterioration model for adults hospitalised with COVID-19: a prospective cohort study. Lancet Respir Med 2021;9:349–59. doi:10.1016/S2213-2600(20)30559-2

34 Hohl CM, Rosychuk RJ, Archambault PM, et al. Derivation and validation of a clinical score to predict death among non-palliative covid-19 patients presenting to emergency departments: the CCEDRRN COVIDmortality score. medRxiv 2021;:2021.07.28.21261283. doi:10.1101/2021.07.28.21261283

35 Halasz G, Sperti M, Villani M, et al. A Machine Learning Approach for Mortality Prediction in COVID-19 Pneumonia: Development and Evaluation of the Piacenza Score. J Med Internet Res 2021;23:e29058–e29058. doi:10.2196/29058

36 Gupta RK, Marks M, Samuels THA, et al. Systematic evaluation and external validation of 22 prognostic models among hospitalised adults with COVID-19: an observational cohort study. Eur Respir J 2020;56:2003498. doi:10.1183/13993003.03498-2020

37 Update to living systematic review on prediction models for diagnosis and prognosis of covid-19. BMJ 2021;372. doi:10.1136/bmj.n236

38 The RECOVERY Collaborative Group. Dexamethasone in Hospitalized Patients with Covid-19. N Engl J Med 2021;384:693–704. doi:10.1056/NEJMoa2021436

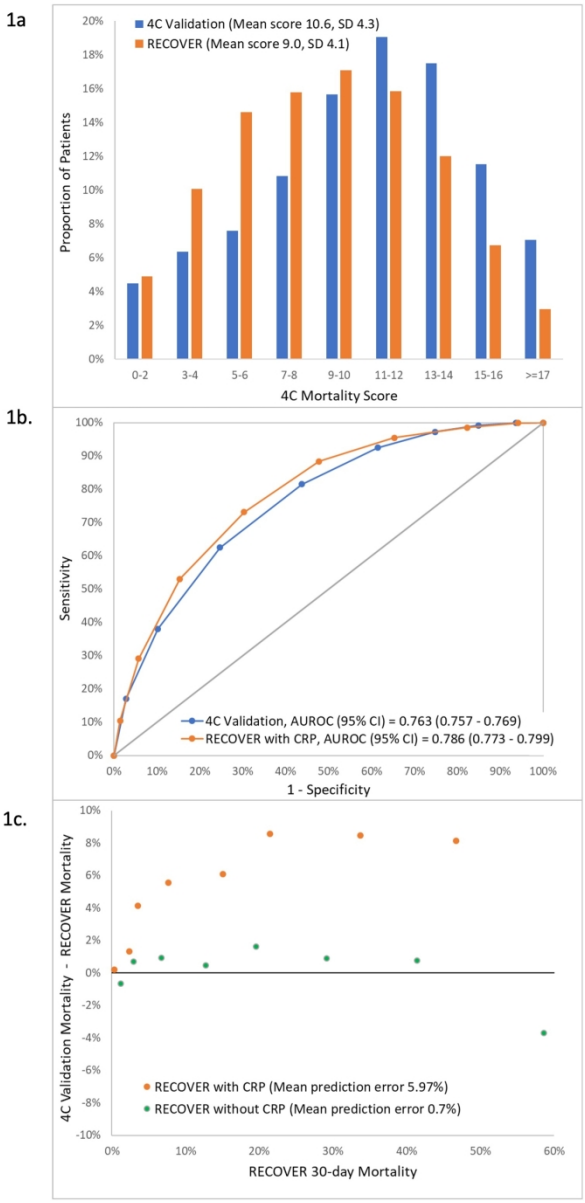

Figure 1: Comparison of 4C validation and RECOVER datasets 1a. 4C Mortality Scores were lower in the RECOVER dataset than in the original 4C validation dataset. 1b. ROC curves for the 4C Mortality Score (categorized into the 9 ranges from Figure 1a) in the 4C validation dataset and the RECOVER dataset. 1c. Calibration plot (modified Bland-Altman) showing prediction error versus observed mortality for the 4C Mortality Score with and without the C-reactive protein (CRP) component. Points from left to right are in the 4C Mortality Score ranges shown in Figure 1A from left to right.

118x238mm (330 x 330 DPI)

Table A. Observed mortality by 4C Risk Group calculated without C-reactive protein for RECOVER dataset of SARS-CoV-2+ patients hospitalized from the Emergency Department.

4C Mortality Risk Group	Patient Distribution	Predicted Risk	Observed Mortality	Prediction Error
0-2	8.1%	0.5%	1.6%	-1.1%
3-4	10.8%	3.7%	3.7%	-0.1%
5-6	14.4%	7.6%	8.4%	-0.7%
7-8	16.9%	13.2%	16.0%	-2.8%
9-10	17.9%	21.2%	25.9%	-4.7%
11-12	16.3%	30.0%	34.4%	-4.4%
13-14	10.3%	42.2%	49.9%	-7.7%
15-16	4.7%	54.9%	62.4%	-7.5%
≥17	0.7%	72.3%	66.1%	6.3%

Mean Prediction Error	-3.3%
√Calibration Error	0.0421
√Brier Score	0.3875

STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No	Recommendation
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract "retrospective cohort study" √ **(page 3)**
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found√ **(page 3)**
Introduction		
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported√**(page 6)**
Objectives	3	State specific objectives, including any prespecified hypotheses√**(page6,7)**
Methods		
Study design	4	Present key elements of study design early in the paper√ **(page6)**
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection **(page 3,7)**
Participants	6	**(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up √(page 6)** *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed *Case-control study*—For matched studies, give matching criteria and the number of controls per case
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable √ **(page 7)**
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group √ **(page 7)**
Bias	9	Describe any efforts to address potential sources of bias √ **(page 3, 12)**
Study size	10	Explain how the study size was arrived at √ **(page 7)**
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why √ **(page 7-9)**
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding √ **(page 8)**
		(*b*) Describe any methods used to examine subgroups and interactions √ **(page 8)**
		(*c*) Explain how missing data were addressed √ **(page 8)**
		(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed *Case-control study*—If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy
		(*e*) Describe any sensitivity analyses

Continued on next page

Results		
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed √
		(b) Give reasons for non-participation at each stage
		(c) Consider use of a flow diagram
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders √
		(b) Indicate number of participants with missing data for each variable of interest
		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount) 30 day mortality √
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time **√ (page 7)**
		Case-control study—Report numbers in each exposure category, or summary measures of exposure
		Cross-sectional study—Report numbers of outcome events or summary measures
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included **√ (page 9,11)**
		(*b*) Report category boundaries when continuous variables were categorized
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses **√ (page 9,10)**
Discussion		
Key results	18	Summarise key results with reference to study objectives**√ (page 9)**
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias**√(page 12)**
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence**√ (page 12)**
Generalisability	21	Discuss the generalisability (external validity) of the study results**√ (page 12)**
Other information		
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based**√ (page 2)**

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.